

# Density matrix of black hole radiation

---

**Lasma Alberte,<sup>a</sup> Ram Brustein,<sup>a</sup> Andrei Khmelnitsky,<sup>a</sup> and A.J.M. Medved<sup>b,c</sup>**

<sup>a</sup> *Department of Physics, Ben-Gurion University, Beer Sheva 84105, Israel*

<sup>b</sup> *Department of Physics & Electronics, Rhodes University, Grahamstown 6140, South Africa*

<sup>c</sup> *National Institute for Theoretical Physics (NITheP), Western Cape 7602, South Africa*

*E-mail:* [lasma@post.bgu.ac.il](mailto:lasma@post.bgu.ac.il), [ramyb@bgu.ac.il](mailto:ramyb@bgu.ac.il),

[andreykh@post.bgu.ac.il](mailto:andreykh@post.bgu.ac.il), [j.medved@ru.ac.za](mailto:j.medved@ru.ac.za)

**ABSTRACT:** Hawking’s model of black hole evaporation is not unitary and leads to a mixed density matrix for the emitted radiation, while the Page model describes a unitary evaporation process in which the density matrix evolves from an almost thermal state to a pure state. We compare a recently proposed model of semiclassical black hole evaporation to the two established models. In particular, we study the density matrix of the outgoing radiation and determine how the magnitude of the off-diagonal corrections differs for the three frameworks. For Hawking’s model, we find power-law corrections to the two-point functions that induce exponentially suppressed corrections to the off-diagonal elements of the full density matrix. This verifies that the Hawking result is correct to all orders in perturbation theory and also allows one to express the full density matrix in terms of the single-particle density matrix. We then consider the semiclassical theory for which the corrections, being non-perturbative from an effective field-theory perspective, are much less suppressed and grow monotonically in time. In this case, the Rényi entropy for the outgoing radiation is shown to grow linearly at early times; but this growth slows down and the entropy eventually starts to decrease at the Page time. In addition to comparing models, we emphasize the distinction between the state of the radiation emitted from a black hole, which is highly quantum, and that of the radiation emitted from a typical classical black body at the same temperature.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Hawking model of black hole evaporation</b>	<b>4</b>
<b>3</b>	<b>The multi-particle density matrix in terms of the single-particle density matrix</b>	<b>6</b>
<b>4</b>	<b>The state of the emitted particles in the Hawking model</b>	<b>7</b>
4.1	A reminder: A scalar field in a box	8
4.2	Emission of localized wave packets	9
4.3	Single-particle density matrix in the wave-packet basis	12
<b>5</b>	<b>Off-diagonal elements of the density matrix</b>	<b>14</b>
5.1	The Hawking model	14
5.2	The Page model	19
5.3	The semiclassical model	21
<b>6</b>	<b>Conclusion</b>	<b>27</b>
<b>A</b>	<b>Expression for the full density matrix in terms of <math>\rho_j^i</math></b>	<b>28</b>
<b>B</b>	<b>Grey-body factors</b>	<b>31</b>
<b>C</b>	<b>Off-diagonal corrections to the Hawking model</b>	<b>32</b>
C.1	Off-diagonal elements in frequency	34
C.2	Off-diagonal elements in mode-occupation number	35
C.3	Final Result	35
<b>D</b>	<b>Higher moments of the Marchenko–Pastur and <math>\rho_{SC}</math> distributions</b>	<b>36</b>

---

## 1 Introduction

There has been a recent spike in activity on understanding the implications of black hole (BH) evaporation [1, 2]. This can be attributed, in large part, to the controversial proposal that a unitary evaporation process comes with the cost of a “firewall” [3] or “energetic curtain” [4]—these being colorful euphemisms for an apparent tension between general relativity and the unitarity of quantum field theory in a BH

setting [5–8]. While the debate rages on, a consensual mechanism for information release is still lacking.

In Hawking’s model, the process of BH evaporation is not unitary [2]. Hawking argued that the correlation functions of the emitted radiation are diagonal in mode-occupancy number, frequency and emission time and, from this, deduced that the density matrix for the radiation is diagonal in the same quantities. In fact, to good accuracy, the final state of the emitted radiation is thermal, and so it is similar to a maximally mixed state. Then the evaporation cannot be the result of unitary evolution from a nearly pure state, which lead Hawking to state (a statement which he more recently retracted [9]) that the process of gravitational collapse is not compatible with the standard principles of quantum mechanics.

The established benchmark model for describing unitary BH evaporation has been the Page model [10]. In this model, the BH and the emitted radiation are assumed to be in a pure state in some large Hilbert space, which is partitioned into an “in” part representing the BH and an “out” part representing the outgoing radiation. Because the combined state is pure, there must be a special basis in which the full density matrix has only a single eigenvalue. An external observer measures the reduced “out” density matrix in some random basis and treats the transformation matrix  $U$  as a random matrix with prescribed statistics [10, 11]. (See also [12] for a recent review.) When the “out” system is smaller than the “in” system, it looks to an external observer as if it were thermal. Conversely, when the “out” system becomes the larger one, the deviations from a thermal state grow and indicate that the total state is indeed pure. The critical time in which the midpoint of evaporation is reached is normally called the Page time. Given that the BH is initially in a pure state and that the evolution is unitary, the Page model can plausibly be viewed as setting the minimal rate for purification of the emitted radiation.

Two of the current authors (RB and AJMM) have recently proposed a semi-classical model of BH evaporation [13, 14], with the premise of repeating Hawking’s seminal calculations [1, 2]—which assume a classical background metric—so as to include a fluctuating BH geometry. The incipient BH is endowed with a quantum wavefunction [15, 16], leading to expectation values in place of fixed classical parameters. From this perspective, Hawking’s model pertains to the limit of an infinitely massive BH with a fixed size, whereas the semiclassical reformulation treats the BH mass as finite with a continually decreasing size due to classical back-reaction effects. In the new model, the correlation functions of the emitted radiation are no longer diagonal and the evaporation process becomes unitary, even though the thermal-like emission spectrum is maintained. This model of unitary BH evaporation obeys the general constraints about the rate of purification. In particular, it has been shown in [17], relying on the results of the current paper, that the Rényi entropy of the radiation decreases monotonically from its peak value at the Page time and does so at a faster rate than the Page model predicts.

From the point of view of an effective field theory in a fixed curved-space background, the correlation functions of the semiclassical model are modified in a non-perturbative way [16]. Contrary to expectations, these modifications could become significant in some situations. It has been shown that, when the leading result in the effective field theory either vanishes or diverges, then the non-perturbative corrections are particularly relevant [18]. In the first case, we expect a correction of the order  $1/S_{BH}$  (the inverse of the BH entropy) to become the leading result and, in the second, the divergence should be replaced by a large but finite  $S_{BH}$ . These small non-perturbative quantum effects could be coherent, and so their amplitudes could add up and grow with time; even growing to the point where they become comparable to the classical outcome [13].

The previous studies of the new model have utilized the single-particle density matrix (the two-point function) of the radiation as the primary tool<sup>1</sup> rather than the full multi-particle density matrix. The single-particle density matrix is easier to calculate and, for the radiation emitted by a BH, sufficient to completely determine the full density matrix. One of the primary goals of the current paper is to make explicit the relationship between the single-particle and multi-particle density matrices. The results are presented in a way that can be applied to the Hawking model, the proposed semiclassical model and even to more general modifications of Hawking’s framework. As an upshot, we are able to find a closed expression for the Rényi entropy of the full density matrix that is expressed directly in terms of the single-particle density matrix.

It has been argued for BHs in Anti-de Sitter (AdS) space—at first by Maldacena [19] with subsequent elaborations by many others — that non-perturbative contributions from other geometries which do not possess a horizon could be relevant to the state of the Hawking radiation. Maldacena’s specific example was the contribution of thermal AdS to the matter correlations for large BHs. In this case, the contribution to the density matrix of the radiation is expected to be exponentially suppressed  $\sim e^{-S_{BH}}$ . However, because the number of off-diagonal elements is exponentially large  $\sim e^{+S_{BH}}$ , one could not be sure as to the relevance of such a contribution. It could then be argued that, without a better control on the non-perturbative contributions from other geometries (which are usually not available), it becomes difficult to trust the Hawking result.

We have two comments on this issue. First, according to our analysis, one requires much larger corrections  $\sim e^{-S_{BH}/2}$  to challenge the reliability of the leading-order outcomes (See, for example, [20]). Second, the quantum fluctuations of the background BH geometry induce off-diagonal effects in frequency space that are only power suppressed  $\sim 1/\sqrt{S_{BH}}$  and sufficient by themselves to restore unitarity. These

---

<sup>1</sup>This matrix was called “the density matrix” by abuse of language or “the radiation matrix” in the previous studies.

statements will be made more precise in due course.

The rest of the paper is organized as follows: In the next section, we review the basic framework of the Hawking model of BH evaporation and recall its main results. Then, in Section 3, we formulate a precise relationship between the multi-particle density matrix and the single-particle density matrix. The preceding analysis is applied to the Hawking model in Section 4, with the aim of regularizing formal divergences that arise in the eternal-BH framework. Section 4 also includes a discussion on how the state of the BH radiation differs from that of thermal radiation emitted by a “normal” black body. The formal analysis up to this point sets the stage for Section 5, where we discuss similarities and differences between the three models of BH evaporation: the Hawking model, the Page model and the semiclassical model. There is then a brief conclusion in Section 6, followed by four appendices that fill in some technical gaps in the main text.

## 2 The Hawking model of black hole evaporation

Let us start here by recalling the original Hawking description of BH evaporation, which completely dismisses the back-reaction and time-dependence effects. This paper uses the notations of Hawking in [2]. In particular, we consider the initial vacuum state  $|0_-\rangle$  and a final vacuum state  $|0_+\rangle$ . The initial vacuum contains no particles with respect to the creation and annihilation operators  $\{a_i^+, a_i\}$  as defined at past null infinity,

$$a_i|0_-\rangle = 0 . \quad (2.1)$$

Following Hawking, the Hilbert space at future infinity is a direct product of the Hilbert space of the particles falling into the BH,  $\mathcal{H}_{\text{in}}$ , and of those going out,  $\mathcal{H}_{\text{out}}$ , so that  $\mathcal{H} = \mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{out}}$ . The ingoing particles are created and annihilated by the operators  $\{c_i^+, c_i\}$ , whereas the outgoing ones are created and annihilated by  $\{b_i^+, b_i\}$ . Both are related to  $\{a_i^+, a_i\}$  via Bogolyubov transformations. The final vacuum state can be represented as  $|0_+\rangle = |0_{\text{in}}\rangle|0_{\text{out}}\rangle$  and contains neither outgoing nor ingoing particles,

$$b_i|0_+\rangle = 0 , \quad c_i|0_+\rangle = 0 . \quad (2.2)$$

The two vacuum states  $|0_-\rangle$  and  $|0_+\rangle$  do not coincide with one another due to the particle creation by the BH. In other words, an initially empty state will appear to have a non-zero occupation number in the final state, meaning that  $b_i|0_-\rangle \neq 0$ . The initial vacuum can be expressed as a linear combination of the Fock states at future infinity as

$$|0_-\rangle = \lambda_{ab}|a_{\text{out}}\rangle|b_{\text{in}}\rangle , \quad (2.3)$$

where  $|a_{\text{out}}\rangle$  and  $|b_{\text{in}}\rangle$  are the basis vectors of the Fock spaces which are spanned by the operators  $\{b_i^+, b_i\}$  and  $\{c_i^+, c_i\}$ , respectively. As a consequence,

$$|a_{\text{out}}\rangle = \prod_j \frac{1}{\sqrt{n_{ja}!}} (b_j^+)^{n_{ja}} |0_{\text{out}}\rangle, \quad |b_{\text{in}}\rangle = \prod_k \frac{1}{\sqrt{n_{kb}!}} (c_k^+)^{n_{kb}} |0_{\text{in}}\rangle. \quad (2.4)$$

The object of interest is the density matrix of the initial vacuum state,

$$\hat{\rho}_{\text{vac}} \equiv |0_{-}\rangle\langle 0_{-}|, \quad (2.5)$$

which allows one to calculate the vacuum expectation values of observables according to the standard rule  $\langle 0_{-} | \hat{\mathcal{O}} | 0_{-} \rangle = \text{tr} \hat{\rho}_{\text{vac}} \hat{\mathcal{O}}$ .

In particular, we will be interested in the expectation values of the observables at future infinity. These will be composed only from the outgoing creation and annihilation operators  $\{b_i^+, b_i\}$  and so can be written as  $\hat{\mathcal{O}} = \hat{\mathcal{O}}_{\text{out}} \otimes \hat{\mathcal{I}}_{\text{in}}$ , where  $\hat{\mathcal{I}}_{\text{in}}$  denotes the identity operator in the Hilbert space of the ingoing states. The expectation value of this operator then takes the form

$$\langle 0_{-} | \hat{\mathcal{O}} | 0_{-} \rangle = \langle 0_{-} | \hat{\mathcal{O}}_{\text{out}} | 0_{-} \rangle = \rho^{\text{out}}_{ac} \mathcal{O}_{\text{out}}^{ca} = \text{tr}_{\text{out}} \hat{\rho}^{\text{out}} \hat{\mathcal{O}}_{\text{out}}. \quad (2.6)$$

Here,  $\rho^{\text{out}}_{ac}$  are the matrix elements of the reduced density matrix of the Hilbert space  $\mathcal{H}_{\text{out}}$  for the outgoing radiation and are determined by

$$\hat{\rho}^{\text{out}} \equiv \text{tr}_{\text{in}} \hat{\rho}_{\text{vac}} = \langle b_{\text{in}} | 0_{-} \rangle \langle 0_{-} | b_{\text{in}} \rangle = \lambda_{ab} \bar{\lambda}_{cb} |a_{\text{out}}\rangle \langle c_{\text{out}}| = \rho^{\text{out}}_{ac} |a_{\text{out}}\rangle \langle c_{\text{out}}|, \quad (2.7)$$

whereas

$$\mathcal{O}_{\text{out}}^{ca} = \langle c_{\text{out}} | \hat{\mathcal{O}}_{\text{out}} | a_{\text{out}} \rangle. \quad (2.8)$$

The elements of the density matrix for the outgoing radiation can be found by calculating different moments of the operators  $\{b_i^+, b_i\}$ . In particular, the number operator of the  $j$ th outgoing mode is

$$\langle n_j \rangle \equiv \langle 0_{-} | b_j^+ b_j | 0_{-} \rangle = \sum_a n_{ja} \rho^{\text{out}}_{aa}. \quad (2.9)$$

The famous result of Hawking is that these elements are the same as for thermal radiation,

$$\rho^{\text{out}}_{ac} = \prod_j \delta_{n_{ja} n_{jc}} P(n_{ja}), \quad P(n) = \frac{(1 - e^{-\omega/T})(e^{-\omega/T}\Gamma)^n}{[1 - (1 - \Gamma)e^{-\omega/T}]^{n+1}}, \quad (2.10)$$

where  $\Gamma$  is the grey-body factor. The density matrix (2.10) is diagonal in both  $j_a$  labeling the mode frequencies and  $n_{ja}$ , the occupation number of each mode. The index  $a$  (also  $c$ ) runs through the set of basis vectors of the Fock space that is built by the operators  $b_i^+$ . Hence, the dimension of the density matrix is formally infinite.

The result of Hawking indicates that the density matrix for the outgoing radiation is thermal and thus is similar to that of a maximally mixed state. This is in contradiction to unitary evolution and at the core of the information paradox since the collapsing matter was in a pure state and evolved into a mixed state. Nonetheless, we will eventually show how the presence of off-diagonal elements in the density matrix  $\hat{\rho}^{\text{out}}$  can lead to the purification of the outgoing radiation in the course of the BH evaporation.

### 3 The multi-particle density matrix in terms of the single-particle density matrix

The goal of this section is to expose the connection of the full density matrix of the outgoing radiation  $\hat{\rho}^{\text{out}}$ , as defined in (2.7), to the widely used single-particle density matrix

$$\rho_j^i = \langle 0_- | b_j^+ b^i | 0_- \rangle . \quad (3.1)$$

Let us begin with the well-known Bogolyubov transformation between the creation and annihilation operators at future infinity  $\{b_i^+, b^i\}$  and at past infinity  $\{a_i^+, a^i\}$ . This is expressible as<sup>2</sup>

$$b^i = \sum_j (\bar{\alpha}_j^i a^j - \bar{\beta}^{ij} a_j^+) , \quad b_i^+ = \sum_j (\alpha_i^j a_j^+ - \beta_{ij} a^j) . \quad (3.2)$$

In terms of the Bogolyubov coefficients, we can express  $\rho_j^i$  as follows:

$$\rho_j^i = \sum_k \bar{\beta}^{ik} \beta_{jk} . \quad (3.3)$$

As shown in Appendix A, there is a closed expression for the density matrix  $\hat{\rho}^{\text{out}}$  in terms of the  $b$ 's,

$$\hat{\rho}^{\text{out}} = \frac{1}{Z} e^{-b_i^+ \Omega^i_j b^j} , \quad (3.4)$$

where  $Z$  is the normalization factor and  $\Omega$  is a  $c$ -number Hermitian matrix. The latter is related to the single-particle density matrix  $\rho_j^i$  as

$$\rho = \frac{1}{e^\Omega - 1} , \quad \text{or} \quad e^\Omega = 1 + \frac{1}{\rho} . \quad (3.5)$$

From now on, the superscript “out” on  $\hat{\rho}$  will be omitted for brevity,

$$\hat{\rho} \equiv \hat{\rho}^{\text{out}} . \quad (3.6)$$

---

<sup>2</sup>Here, upper and lower indices have been introduced so as to make the matrix multiplication explicit. The summation symbol will typically be omitted in what follows.

The full density matrix  $\hat{\rho}$  is thus completely defined by  $\rho_j^i = \bar{\beta}^{ik}\beta_{jk}$ , with the normalization factor  $Z$  expressed as

$$Z = \det \left[ \frac{1}{1 - e^{-\Omega}} \right] = \det [1 + \rho] . \quad (3.7)$$

One can define the entropy and related quantities directly in terms of the matrix  $\rho$ . In particular, an explicit expression for the Rényi entropy,  $H_2 \equiv -\ln \frac{\text{tr} [\hat{\rho}^2]}{(\text{tr} \hat{\rho})^2}$ , is readily obtained by evaluating

$$\begin{aligned} \text{tr} [\hat{\rho}^2] &= \frac{1}{Z^2} \text{tr} [e^{-2b^+\Omega b}] = \det [1 - e^{-\Omega}]^2 \det \left[ \frac{1}{1 - e^{-2\Omega}} \right] = \\ &= \det \left[ \frac{1 - e^{-\Omega}}{1 + e^{-\Omega}} \right] = \det \left[ \frac{1}{1 + 2\rho} \right] , \end{aligned} \quad (3.8)$$

from which it follows that

$$H_2 = \text{tr} [\ln (1 + 2\rho)] . \quad (3.9)$$

From this expression, it would appear that the Rényi entropy does not vanish unless  $\rho = 0$ , when no particles have been emitted. However, the entropy for a state with  $N$  radiated particles can reach as low as  $H_2 = \ln (2N) \ll S_{BH}(0)$ , which is below the scale of validity of the approximations that were used in deriving Eq. (3.9).

The expression for the von Neumann entropy is more cumbersome but we include it for completeness:

$$S = -\text{tr} [\hat{\rho} \ln \hat{\rho}] = \text{tr} \ln [1 + \rho] + \det \left[ \rho \ln \left( 1 + \frac{1}{\rho} \right) \right] . \quad (3.10)$$

In this paper, we will therefore concentrate on the Rényi entropy  $H_2$ . It is much easier to calculate and provides an accurate measure of the entanglement in the state of the BH radiation.<sup>3</sup>

## 4 The state of the emitted particles in the Hawking model

Having a formal expression for the density matrix at hand, we can now evaluate the Rényi entropy and the particle number in the state of the outgoing radiation in the Hawking framework. However, as made clear below, these quantities are formally proportional to an infinite sum over all the possible modes in the Fock space of the outgoing radiation. This divergence comes about from implicitly assuming an eternal BH that radiates for an infinite period of time and, hence, emits an infinite amount of particles.

---

<sup>3</sup>For Gaussian states, such as the vacuum state of a non-interacting theory,  $H_2$  respects the strong-subadditivity condition in the same way the von Neumann entropy does [21].

The physical situation, however, must be different: The number of particles emitted by the BH during a finite time interval is finite and depends on the initial mass of the BH. The goal of this section is to present a way of counting the modes emitted by the BH which reach an observer at future infinity during some finite time interval.<sup>4</sup> By implementing this method, we find that the physical quantities indeed become finite. For simplicity, the analysis is initially specialized to the case of the Hawking model.

In the current case, the matrix  $\Omega^i_j = \frac{\omega_i}{T} \delta^i_j$  is diagonal, and Eqs. (3.7) and (3.9) for the partition function and the Rényi entropy then read

$$-\ln Z = \sum_i \ln (1 - e^{-\omega_i/T}) , \quad (4.1)$$

$$H_2 = \sum_i \ln \left( \frac{1 + e^{-\omega_i/T}}{1 - e^{-\omega_i/T}} \right) , \quad (4.2)$$

while the total particle number is given by

$$N = \sum_i \frac{1}{e^{\omega_i/T} - 1} . \quad (4.3)$$

The sums in the above expressions run over the full set of the frequency modes  $\omega_i$  of a quantum field in the BH background. Hence, the dimension of the set of the possible frequencies is infinite, leading to an infinite entropy and particle number.

#### 4.1 A reminder: A scalar field in a box

To better understand the correct way of counting the modes so as to render physical quantities finite, let us first recall a simple case: A massless quantum field in a box.

A massless quantum scalar field  $\phi(t, \vec{x})$  in a box of size  $L$  with reflecting boundary conditions (*i.e.*,  $\phi(t, 0) = \phi(t, L) = 0$ ) can be expanded in terms of the Fourier modes

$$u_{\vec{k}}(t, \vec{x}) \equiv \frac{1}{(2\pi L)^{3/2}} \frac{1}{(2\omega)^{1/2}} \sin \left( \frac{2\pi}{L} k_x x \right) \sin \left( \frac{2\pi}{L} k_y y \right) \sin \left( \frac{2\pi}{L} k_z z \right) e^{-i\omega t} , \quad (4.4)$$

with frequencies  $\omega = \frac{2\pi}{L} \sqrt{k_x^2 + k_y^2 + k_z^2}$ , where  $k_x$ ,  $k_y$  and  $k_z$  can be any integer.

For a large enough box ( $L \gg 1/T$ ), the sum over the modes can be approximated by an integral over a continuous momentum  $\vec{p}$ ,

$$\sum_{k_x, k_y, k_z} \simeq L^3 \int \frac{d^3 p}{(2\pi)^3} , \quad (4.5)$$

with the corresponding frequency given by  $\omega(\vec{p}) = |\vec{p}|$ . In this way, the quantities in Eqs. (4.1)–(4.3) are proportional to the volume of the box  $V = L^3$ ; that is, they are

---

<sup>4</sup>For an earlier, detailed calculation of the physical emission rates that employs a different technique, see Ref. [22].

extensive. Moreover, since the only other dimensional parameter is the temperature  $T$ , any extensive, dimensionless quantity is given by  $VT^3$  times some numerical constant:

$$-\ln Z = VT^3 \cdot \frac{1}{2\pi^2} \int_0^\infty dx x^2 \ln(1 - e^{-x}) = -\frac{\pi^2}{90} VT^3, \quad (4.6)$$

$$H_2 = VT^3 \cdot \frac{1}{2\pi^2} \int_0^\infty dx x^2 \ln\left(\frac{1 + e^{-x}}{1 - e^{-x}}\right) = \frac{\pi^2}{48} VT^3, \quad (4.7)$$

$$N = VT^3 \cdot \frac{1}{2\pi^2} \int_0^\infty dx \frac{x^2}{e^x - 1} = \frac{\zeta(3)}{\pi^2} VT^3, \quad (4.8)$$

where  $x = \omega/T$ .

In general, due to the diagonal form of the thermal density matrix, any quantity that is given by a trace over the Hilbert space  $\mathcal{H}_{\text{out}}$  is proportional to a product over the available frequencies (as in Eq. (A.7)). A logarithm of such a quantity is then proportional to a sum over the frequencies and is extensive by virtue of Eq. (4.5). Imposing some physical boundary conditions on the scalar field results in a restriction on the allowed frequency modes—in this case, the condition that the wave numbers have to be integer multiples of  $2\pi/L$ . The formally infinite quantities (4.1) - (4.3) then become finite, as should be true for any physical system.

## 4.2 Emission of localized wave packets

We have just seen that, for a field in a box, the frequencies are quantized in units of the inverse size of the box and the total number of modes is proportional to the volume of the box. In order to study how a BH radiates into an infinite space, it is more convenient to use localized wave packets with finite normalization instead of the Fourier modes (as in [1]).

Let us begin here with a scalar field that is propagating in the background of a BH. It can be expanded in terms of the complete set of solutions of the wave equation having only positive frequencies. At past infinity, the expansion contains only ingoing modes and takes the form

$$\phi = \sum_i (f_i a_i + \bar{f}_i a_i^+) , \quad (4.9)$$

where  $\{a_i^+, a_i\}$  are the creation and annihilation operators defined at past infinity and the sum runs over a discrete set of modes  $\{f_i\}$  with finite normalization. At future infinity, the expansion of the scalar field in terms of positive-frequency solutions contains both ingoing and outgoing modes,

$$\phi = \phi_{\text{in}} + \phi_{\text{out}}, \quad \phi_{\text{out}} = \sum_i (p_i b_i + \bar{p}_i b_i^+) , \quad (4.10)$$

where  $\{b_i^+, b_i\}$  are the creation and annihilation operators of the outgoing modes  $\{p_i\}$  at future infinity.

In the continuum normalization with frequencies  $\omega$  (rather than the discrete wave packets  $i$ ), the ingoing and outgoing Fourier modes for the spherically symmetric solutions are expressible as

$$f_{\omega lm}(v, r, \theta, \phi) = F_{\omega lm}(r) Y_l^m(\theta, \phi) e^{i\omega v}, \quad (4.11)$$

$$p_{\omega lm}(u, r, \theta, \phi) = P_{\omega lm}(r) Y_l^m(\theta, \phi) e^{i\omega u}, \quad (4.12)$$

where  $l$  and  $m$  are the angular-momentum numbers, and  $v$  and  $u$  are the advanced and retarded tortoise coordinates (respectively).

A complete set of localized wave packets can then be defined as [1]

$$f_{jnlm}(v, r, \theta, \phi) = \varepsilon^{-1/2} \int_{j\varepsilon}^{(j+1)\varepsilon} e^{-2\pi i n \omega / \varepsilon} f_{\omega lm}(v, r, \theta, \phi) d\omega, \quad (4.13)$$

$$p_{jnlm}(u, r, \theta, \phi) = \varepsilon^{-1/2} \int_{j\varepsilon}^{(j+1)\varepsilon} e^{-2\pi i n \omega / \varepsilon} p_{\omega lm}(u, r, \theta, \phi) d\omega. \quad (4.14)$$

In place of the continuous label  $\omega$ , the wave packets have been labeled by two integer indices  $j \geq 0$  and  $n$ . The wave packet with index  $j$  contains waves with frequencies that are localized in the range  $j\varepsilon \leq \omega \leq (j+1)\varepsilon$ . Index  $n$  labels the ray along which the packet is propagating: The wave packet  $f_{jn}$  is peaked around the advanced time  $v = 2\pi n \varepsilon^{-1}$  and the wave packet  $p_{jn}$ , around the retarded time  $u = 2\pi n \varepsilon^{-1}$ , both having a width of  $2\pi\varepsilon^{-1}$ .

Our particular interest is the wave packets which were emitted by the collapsing body and then detected at some fixed distance away from the BH during a given period of time  $t_0 \leq t \leq t_0 + \Delta t$ . Such wave packets will also be localized in the corresponding range of retarded time  $u_0 \leq u \leq u_0 + \Delta t$ . For a long enough time interval, one can always choose the width  $\varepsilon^{-1}$  so as to ensure the detection of many wave packets ( $\varepsilon^{-1} \ll \Delta t$ ), as well as maintain a fine-enough frequency resolution ( $\varepsilon \ll T$ ) for each wave packet to be treated as a monochromatic mode of fixed frequency. In such a case, the summation over the wave-packet position  $n$  in the interval  $\Delta n = \Delta t / (2\pi\varepsilon^{-1})$  can be approximated as

$$\sum_{n=n_0}^{n_0+\Delta n} \approx \frac{\varepsilon}{2\pi} \Delta t, \quad (4.15)$$

whereas the summation over the discrete set of frequencies  $j = \omega / \varepsilon$  can be approximated by the integral

$$\sum_{j=0}^{\infty} \approx \int_0^{\infty} \frac{d\omega}{\varepsilon}. \quad (4.16)$$

As one can now see, in this approximation, the total number of modes which can be detected during the time interval  $\Delta t$ , does not depend on the choice of the

parameter  $\varepsilon$  and is proportional to  $\Delta t$  times the integral over mode frequencies,

$$\sum_i = \sum_{j=0}^{\infty} \sum_{n=n_0}^{n_0+\Delta n} = \Delta t \int_0^{\infty} \frac{d\omega}{2\pi}. \quad (4.17)$$

The total number of particles detected during the time interval  $\Delta t$  in this case is given by

$$N = \sum_{j=0}^{\infty} \sum_{n=n_0}^{n_0+\Delta n} \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{\Gamma_{jlm}}{e^{j\varepsilon/T} - 1} = \Delta t \sum_{l=0}^{\infty} \sum_{m=-l}^l \int_0^{\infty} \frac{d\omega}{2\pi} \frac{\Gamma_{\omega lm}}{e^{\omega/T} - 1}, \quad (4.18)$$

where  $\Gamma_{\omega lm}$ , the so-called grey-body factor, is determined by the properties of the modes near the horizon.

The emission of the modes with high multipoles is highly suppressed,  $\Gamma_{\omega lm} \ll \Gamma_{\omega 00} \equiv \Gamma(\omega)$  for  $l, m > 0$ , and the sum (4.18) can be well approximated with the  $l = m = 0$  term only (see, *e.g.*, [22]). In what follows, we will consider only the  $l = m = 0$  modes and omit the angular-momentum labels. Moreover, in the Schwarzschild case, there is only one available dimensional parameter, which can be chosen as the BH temperature  $T$ . Therefore, the grey-body factor  $\Gamma(\omega)$  depends on the frequency only through the ratio  $x = \omega/T$ , and the number of emitted particles simplifies as follows:

$$N = \Delta t T \int_0^{\infty} \frac{dx}{2\pi} \frac{\Gamma(x)}{e^x - 1} \sim \Delta t A T^3 \int_0^{\infty} \frac{dx}{2\pi} \frac{\Gamma(x)}{e^x - 1}, \quad (4.19)$$

where  $A \sim T^{-2}$  is the area of the BH horizon.

Thus, the rate of BH emission coincides with the thermal emission rate of a body with area  $A$ . In reality, due to the non-trivial frequency dependence of the grey-body factors, the spectrum for BH radiation is quite different from the thermal spectrum of a body in an empty box. Some details about the grey-body factors are provided in Appendix B.

And so, in the absence of other dimensional parameters, all the extensive quantities describing the emitted radiation are proportional to the product  $\Delta t T$  or, equivalently, to the total number of emitted particles (4.19). For example, the Rényi entropy  $H_2$  is given by

$$H_2 = \Delta t \int_0^{\infty} \frac{d\omega}{2\pi} \ln \left( 1 + \frac{2\Gamma(\omega/T)}{e^{\omega/T} - 1} \right) = \Delta t T \int_0^{\infty} \frac{dx}{2\pi} \ln \left( 1 + \frac{2\Gamma(x)}{e^x - 1} \right) \propto N. \quad (4.20)$$

Hence, by properly accounting for the finite duration of detection, one finds that the number of particles emitted by the BH, as well as the entropy of the radiation, becomes finite as opposed to the idealized case of infinite emission time.

### 4.3 Single-particle density matrix in the wave-packet basis

In order to highlight the connection between the localized wave-packet description and the semiclassical model as presented in [13, 14], we shall construct the single-particle density matrix  $\rho^i_j$  in the finite wave-packet basis.

Using the definition of the wave packets (4.14), one can express the matrix elements of  $\rho$  in terms of the same Fourier-mode basis,

$$\rho^{jn}_{j'n'} = \varepsilon^{-1} \int_{j\varepsilon}^{(j+1)\varepsilon} d\omega e^{2\pi i n \omega / \varepsilon} \int_{j'\varepsilon}^{(j'+1)\varepsilon} d\omega' e^{-2\pi i n' \omega' / \varepsilon} \rho(\omega, \omega') , \quad (4.21)$$

where

$$\rho(\omega, \omega') = \int_0^\infty d\tilde{\omega} \bar{\beta}_{\omega\tilde{\omega}} \beta_{\omega'\tilde{\omega}} . \quad (4.22)$$

Each wave packet is labeled by the characteristic frequency  $\omega_j = j\varepsilon$  and the detection time  $t_n = 2\pi\varepsilon^{-1}n$ . The choice of  $\varepsilon$  parametrizes a trade-off between the frequency and the position resolutions. The total number of modes  $W$ , within a given range of frequencies  $\Delta\omega$  and an interval of time  $\Delta t$ , does not depend on  $\varepsilon$  and is just given by  $W = \Delta j \cdot \Delta n = \Delta t \Delta\omega / (2\pi)$ . In the case of a thermal distribution, the frequency range  $\Delta\omega$  is effectively set by the temperature. Hence, the total number of modes with substantial occupation numbers is of the order  $W = T \Delta t$ . The matrix  $\rho^{jn}_{j'n'}$  in Eq. (4.21) is, therefore, effectively a  $W \times W$  matrix.

An important feature of BH radiation is that the emission rate  $\Gamma$  is the same as the temperature  $T$ . For a general radiating body, this is not true; even if  $\Gamma \propto T$ , the proportionality constant need not be close to unity. This has important consequences, which we now elaborate on.

From the above discussion on wave packets, we have learned that the total number of emitted modes is given by  $W = \Delta\omega\Delta t$ , where  $\Delta\omega$  is the range of emitted frequencies and  $\Delta t$  is the detection time. For the case of a nearly thermal emitter, the frequency range is given by the temperature  $\Delta\omega \sim T$ , so that  $W = T\Delta t$ . On the other hand, the total number of emitted particles during this same time is given by  $N = \Gamma\Delta t$ , where  $\Gamma$  is the emission rate. For a BH, with the emission rate  $\Gamma \sim T$ , the number of emitted particles  $N$  is then of the same order as the total number of occupied frequency modes  $W$ ; *i.e.*,  $N \sim W$ . This fact has two important consequences for BH radiation:

First, the average occupation number for the modes of radiation is approximately of order unity, as  $N/W \sim \Gamma/T \sim 1$ ; meaning that, on average, each mode of radiation is occupied by only a few particles. Second, since  $W \sim N$ , the single-particle density matrix  $\rho$  is effectively of size  $N \times N$ . Then, as the average occupation numbers correspond to the diagonal elements of the matrix  $\rho$ , these elements go as  $\rho^{jn}_{jn} \sim 1$ . Hence, for BH radiation, the two-point function  $\rho$  in the wave packet basis has the form of an  $N \times N$  *identity* matrix.

The fact that, on average, each mode is occupied by only a few particles also implies that the radiation field in a BH background cannot be treated as semiclassical. This is not the normal state of affairs. For any other thermal emitter, the radiation rate is proportional to its area,  $\Gamma \sim A \cdot T^3$ , and the area for classical emitters is much larger than the inverse temperature squared (*i.e.*, the square of the typical wavelength), so that  $\Gamma/T \sim AT^2 \gg 1$ . Consequently,  $N/W \sim \Gamma/T \gg 1$  and the radiation field is highly classical.

In order to make these ideas more precise, let us calculate  $\text{tr } \rho$  and  $\text{tr } \rho^2$  for the Hawking model [1, 2]. For a large BH,

$$\rho(\omega, \omega') = \delta(\omega - \omega') \rho(\omega) \equiv \delta(\omega - \omega') \frac{\Gamma(\omega)}{e^{\omega/T} - 1}. \quad (4.23)$$

With this expression, in the wave-packet basis,

$$\rho^{jn}{}_{j'n'} = \varepsilon^{-1} \delta_{j'}^j \int_{j\varepsilon}^{(j+1)\varepsilon} d\omega e^{2\pi i (n-n')\omega/\varepsilon} \rho(\omega). \quad (4.24)$$

As one can observe,  $\rho^{jn}{}_{j'n'}$  depends only on the difference in detection times  $t_n - t_{n'}$ . Because of the finite time and frequency resolution, it has non-zero off-diagonal elements with  $n \neq n'$ , which are concentrated in a narrow strip  $|n - n'| \lesssim \varepsilon/T$  corresponding to  $|t_n - t_{n'}| \lesssim T^{-1}$ . The number of non-zero elements of  $\rho$  is, therefore, proportional to its dimension  $W$  and, consequently, of the order of the number of emitted particles  $N$ .

The traces of  $\rho$  and  $\rho^2$  can be explicitly calculated:

$$\text{tr } \rho = \sum_{j=0}^{\infty} \sum_{n=n_0+1}^{n_0+\Delta n} \rho^{jn}{}_{jn} = \varepsilon^{-1} \sum_{n=n_0+1}^{n_0+\Delta n} \int_0^{\infty} d\omega \rho(\omega) = \Delta t \int_0^{\infty} \frac{d\omega}{2\pi} \rho(\omega) \sim \Delta t T = W \sim N, \quad (4.25)$$

$$\begin{aligned} \text{tr } \rho^2 &= \sum_{j=0}^{\infty} \sum_{\substack{n=n_0+1 \\ j'=0 \\ n'=n_0+1}}^{n_0+\Delta n} \rho^{jn}{}_{j'n'} \rho^{j'n'}{}_{jn} \\ &= \varepsilon^{-2} \sum_{j=0}^{\infty} \sum_{\substack{n=n_0+1 \\ n'=n_0+1}}^{n_0+\Delta n} \int_{j\varepsilon}^{(j+1)\varepsilon} d\omega \int_{j\varepsilon}^{(j+1)\varepsilon} d\omega' e^{2\pi i (n-n')(\omega-\omega')/\varepsilon} \rho(\omega) \rho(\omega') \\ &= \varepsilon^{-2} \sum_{j=0}^{\infty} \int_{j\varepsilon}^{(j+1)\varepsilon} d\omega \int_{j\varepsilon}^{(j+1)\varepsilon} d\omega' \left( \frac{\sin [\Delta n \pi (\omega - \omega')/\varepsilon]}{\sin [\pi (\omega - \omega')/\varepsilon]} \right)^2 \rho(\omega) \rho(\omega') \\ &\approx \Delta t \int_0^{\infty} \frac{d\omega}{2\pi} \rho(\omega)^2 \sim \Delta t T \sim N. \end{aligned} \quad (4.26)$$

In order to obtain the last line, we have used that  $(\sin(\Delta n \pi x)/\sin(\pi x))^2 \approx \Delta n \delta(x)$  for large  $\Delta n$ .

The number of the non-zero eigenvalues, which is just the number of diagonal elements in the Hawking model, can be then estimated by the participation ratio (*i.e.*, the inverse of the purity; see, *e.g.*, [23]),

$$PR \equiv \frac{(\text{tr} [\rho])^2}{\text{tr} [\rho^2]} \sim N . \quad (4.27)$$

The result grows linearly with the number of emitted particles and also provides an estimate of the Rényi entropy (4.20).

For a more general case including off-diagonal elements, one can still expect  $\text{tr} \rho \sim N$ , as this result only depends on the dimensionality of the matrix. On the other hand,  $\text{tr} \rho^2$  can be much different, but its value can be readily estimated if the number of non-zero eigenvalues is known or *vice versa*.

## 5 Off-diagonal elements of the density matrix

With the previous framework in hand, we will now consider the structure of the off-diagonal elements in both the single- and multi-particle density matrices. This will be carried out for the three distinct models of an evaporating BH: (1) Hawking's model [1], (2) the Page model [10] and (3) the semiclassical model [13, 14].<sup>5</sup>

In this section, all dimensional quantities are expressed in units of the BH temperature,  $N = N(t)$  is the total number of emitted particles at time  $t$  and  $S_{BH} \equiv S_{BH}(N) = S_{BH}(0) - N(t)$  is the BH entropy at “time”  $N$ . This relation is modified, semiclassically, by fluctuations in the emission time of the Hawking modes. However, the fluctuations were shown to be small [13], being proportional to inverse powers of  $S_{BH}$ .

One of the main issues to be clarified is to what extent the off-diagonal elements in terms of mode-occupation number of the full density matrix are suppressed. If they are indeed highly suppressed, then the relationship between the density matrix  $\hat{\rho}$  and the single-particle density matrix  $\rho$  is reliable.

### 5.1 The Hawking model

In Hawking's model, the outgoing radiation is described to a very good approximation by the diagonal density matrix given in Eq. (2.10). What we would like to address here is how accurate an approximation this is by quantifying the off-diagonal contribution.<sup>6</sup> We shall, however, not discuss any back-reaction nor evolution effects

---

<sup>5</sup>For the Page model, the single-particle density matrix cannot be calculated without additional information because this framework is based upon choosing an arbitrary random basis in the combined radiation and BH Hilbert space. How this basis is related to the Fock space of the radiation is left unspecified.

<sup>6</sup>By off-diagonal contribution, we mean that beyond the broadening effects of the wave-packet treatment in the last section.

to the Hawking model in this subsection. These effects were originally disregarded by Hawking himself and introducing them would mean deviating from the original model in a significant way. We do account for the evolution effects in the semiclassical model that we discuss in the subsection 5.3.

The Hawking density matrix  $\hat{\rho}_H$  is only approximately diagonal in two obvious ways: Elements between states differing in total occupation number are only approximately vanishing, as are those between states with some modes differing in frequency or in the time of emission. We would like to estimate the magnitude of the off-diagonal elements of the density matrix. As far as we know, this issue has not been addressed directly in the literature, although Hawking did expect such corrections to be exponentially small [1].

What we find below is an overall suppression for off-diagonal elements in mode-occupation number that is stronger than  $e^{-S_{BH}}$ . This is significant because it implies that the results of Hawking are not modified in a meaningful way by off-diagonal corrections.<sup>7</sup> We also find that the off-diagonal elements of the single-particle density matrix  $\rho_{\omega}^i$  are smaller than  $1/S_{BH}$ , leading to corrections to physical quantities, such as the entropy, that are suppressed by at least a factor of  $1/S_{BH}^2$ . These are negligible at all times, as originally anticipated by Hawking. Corrections to the diagonal elements of  $\rho_{\omega}^i$  are similarly suppressed by at least  $1/S_{BH}$  and readily absorbed into the normalization.

The fact that the Hawking density matrix is diagonal can be understood by looking at Hawking's derivation of the single-particle density matrix  $\rho_H$ , as defined in Eq. (4.22). The relevant factor goes as  $\rho_H(\omega, \omega') \sim I_-(\omega, \omega')$  such that [1]

$$I_-(\omega, \omega') = \int_0^\infty d\tilde{\omega} (\tilde{\omega})^{-1+i(\omega-\omega')} = \int_{-\infty}^\infty dy e^{+iy(\omega-\omega')} = 2\pi \delta(\omega - \omega') , \quad (5.1)$$

where  $y = \ln \tilde{\omega}$  has been used. This equation makes it clear that the matrix elements of  $\rho_H$  are diagonal in frequency space.

The off-diagonal elements of  $\hat{\rho}_H$  for different occupation numbers are also vanishing, as follows from an estimation of the products

$$\alpha\beta \equiv \int_0^\infty d\tilde{\omega} \alpha_{\omega\tilde{\omega}} \beta_{\omega'\tilde{\omega}} , \quad \bar{\alpha}\bar{\beta} \equiv \int_0^\infty d\tilde{\omega} \bar{\alpha}_{\omega\tilde{\omega}} \bar{\beta}_{\omega'\tilde{\omega}} . \quad (5.2)$$

The products  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  arise in the expectation values of operators like

$$\langle 0_- | (b^+)^n b^{n+2k} | 0_- \rangle \sim (\beta\bar{\beta})^n \cdot (\alpha\beta)^k , \quad (5.3)$$

$$\langle 0_- | (b^+)^{n+2k} b^n | 0_- \rangle \sim (\beta\bar{\beta})^n \cdot (\bar{\alpha}\bar{\beta})^k \quad (5.4)$$

---

<sup>7</sup>See, however, the discussion in the Introduction about corrections arising from geometries without a horizon.

and, thus, define the elements of the density matrix between states that differ in total occupation number by  $\Delta\mathcal{N} = 2k$ . The products  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  also appear in the computation of the generating function (A.2). Consequently, the dependence of the density matrix on  $\rho$  is only truly valid in the case when the products  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  are negligible.

In Hawking's Fourier-space analysis, such terms contain an integral that is similar to the one in Eq. (5.1) but now the two frequencies appear with the same sign,

$$I_+(\omega, \omega') = \int_0^\infty d\tilde{\omega} (\tilde{\omega})^{-1+i(\omega+\omega')} = \int_{-\infty}^\infty dy e^{iy(\omega+\omega')} = 2\pi \delta(\omega + \omega') . \quad (5.5)$$

Since both frequencies are positive, the argument of the delta function can only vanish if both of the frequencies vanish. This leads to one or more factors of  $\delta(\omega + \omega')$  when contracting pairs of operators in the corresponding density-matrix element, as there would necessarily be unequal numbers of creation and annihilation operators within the expectation value.

To help reveal where corrections to off-diagonal elements can appear, let us consider the integration limits in Eqs. (5.1) and (5.5). These limits are an idealization, as a frequency can never be truly infinite nor exactly zero in physically realistic situations. For instance, the upper limit on  $\tilde{\omega}$  should be an exponentially large number, representing ultra-high frequencies on past null infinity that need to be red-shifted in order to produce the thermal radiation at future null infinity. An estimate of this upper limit in units of the BH mass  $M$  is  $\sim e^{\tau/M} \sim e^{(M/M_{Pl})^2} \sim e^{S_{BH}(0)}$ , where  $\tau \sim M^3$  denotes the lifetime of the BH (see, *e.g.*, [24]). As for the lower limit on  $\tilde{\omega}$ , a natural choice is to fix it as the inverse of the BH lifetime  $(\tau/M)^{-1} \sim S_{BH}^{-1}(0)$ . However, in Hawking's model, the BH is regarded as eternal, and so this limit should rather be the exponential of a large negative number.

In the following, we replace the idealized limits with suitable ultraviolet and infrared cutoffs,  $e^{y_{max}}$  and  $e^{-y_{min}}$ , which are assumed to be of order  $y_{max} \sim y_{min} \sim S_{BH}(0)$ . We will further set  $y_* \equiv y_{max} = y_{min}$ , as this symmetry will simplify the calculations without affecting the conclusions. This estimate determines the strength of the off-diagonal corrections of the density matrix in mode-occupation number. If, for some reason, the limits are such that  $y_{max} \sim y_{min} \ll S_{BH}(0)$ , then the original Hawking calculation is significantly modified and the original conclusion about the nature of the density matrix needs to be revised.

The new choice of limits leads to a correction to the previous delta function,

$$\int_{-y_*}^{y_*} dy e^{iy(\omega \pm \omega')} = 2y_* \frac{\sin(y_*(\omega \pm \omega'))}{y_*(\omega \pm \omega')} . \quad (5.6)$$

The tails of this “regulated” delta function provide the off-diagonal correction of interest. Notice that this comes about for both  $I_{\pm}(\omega, \omega')$ , and, hence, the off-diagonal

elements of the density matrix can now be non-vanishing, even when connecting states that differ in total occupation number.

To better understand the implications of these off-diagonal corrections, the single-particle density matrix (and its related products) can be convolved with the same wave packets as in Eq. (4.21). This analysis is carried out in detail in Appendix C, from which there are two main conclusions: First, we find that the off-diagonal elements of the *single-particle* density matrix  $\rho^{jn}_{j'n'} \equiv (\bar{\beta}\beta)^{jn}_{j'n'}$  are suppressed relative to the diagonal elements by a factor of  $\mu_*^{-2} J^{-1}$ , where  $\mu_* \equiv y_*\varepsilon/(4\pi^2 T)$  and  $J = j' - j$ . Second, it is found that all the elements of the products  $\alpha\beta, \bar{\alpha}\bar{\beta}$  are subleading and suppressed as  $\mu_*^{-2} J_{\alpha\beta}^{-1}$ . Here,  $J_{\alpha\beta} = j' + j + 1$  and can be treated as some typical average value of the frequency label  $j$  referring to the frequency range  $\omega_j \in [j\varepsilon, (j+1)\varepsilon]$ . These products set the order of magnitude of the off-diagonal elements of the *full* density matrix. In particular, this result tells us that the leading-order contribution to a typical off-diagonal element of the density matrix<sup>8</sup> is relatively suppressed by a factor of  $\mu_*^{-|\Delta\mathcal{N}|} j_*^{-\frac{1}{2}|\Delta\mathcal{N}|}$ . Here  $\Delta\mathcal{N}$  is the difference in total occupation number of the corresponding states and  $j_*$  can be viewed as a typical value for the integer  $j$ .

Let us elaborate on this last claim by, first, understanding the structure of the density matrix  $\hat{\rho}$ . A natural way of organizing the elements of the density matrix which connect different Fock states is to classify them by the total number of particles in each state. A density-matrix element  $\hat{\rho}_{\mathcal{N}\mathcal{N}'}^{\mathcal{N}}$  then denotes a block of elements connecting states with total occupation numbers  $\mathcal{N}$  and  $\mathcal{N}'$ . Thus, each element  $\hat{\rho}_{\mathcal{N}\mathcal{N}'}^{\mathcal{N}}$  is a matrix of dimension  $d_{\mathcal{N}'} \times d_{\mathcal{N}}$ , where  $d_{\mathcal{N}}$  is given by the size of the Fock subspace spanned by the states with total occupation number equal to  $\mathcal{N}$ . If we denote the number of available frequency modes by  $W$ , then  $d_{\mathcal{N}}$  coincides with the number of distinct elements in a symmetric tensor of rank  $\mathcal{N}$  in a  $W$ -dimensional vector space:

$$d_{\mathcal{N}} = \binom{\mathcal{N} + W - 1}{\mathcal{N}}. \quad (5.7)$$

Suppose that our interest is the strength of the density-matrix element  $\hat{\rho}_{\mathcal{N}\mathcal{N}'}^{\mathcal{N}}$ , such that  $\Delta\mathcal{N} = \mathcal{N} - \mathcal{N}'$  is an even, non-zero integer. (There are no contributions when this difference is odd; *cf.* Eqs. (5.3) and (5.4).) The strength of such an element is determined by a higher-order moment than the two-point function; rather, by an  $(\mathcal{N} + \mathcal{N}')$ -point function with  $\mathcal{N}$  annihilation operators and  $\mathcal{N}'$  creation operators. But, because the theory of interest is non-interacting, the relative strength of such an element can be assessed by looking at products of two-point functions  $\bar{\beta}\beta, \alpha\beta, \bar{\alpha}\bar{\beta}$ . Here, we will be able to correctly match a creation operator with an annihilation operator for all but  $\frac{|\Delta\mathcal{N}|}{2}$  of the total number of operator pairs  $\frac{\mathcal{N} + \mathcal{N}'}{2}$ . Each mismatched

---

<sup>8</sup>Typical in the sense that only an exponentially small fraction of the off-diagonal elements deviate from this behavior.

pair—two creation operators or two annihilation operators—will then contribute one suppression factor of  $\mu_*^{-2} j_*^{-1}$ .

Formally, the density matrix can be represented in terms of the products  $\beta\bar{\beta}$ ,  $\alpha\beta$ , and  $\bar{\alpha}\bar{\beta}$  as

$$\begin{aligned}
\mathcal{N}' = 1 &\rightarrow \begin{pmatrix} \beta\bar{\beta} & 0 & \bar{\alpha}\bar{\beta} & 0 & (\bar{\alpha}\bar{\beta})^2 & 0 & \dots \end{pmatrix} \\
\mathcal{N}' = 2 &\rightarrow \begin{pmatrix} 0 & \beta\bar{\beta} & 0 & \bar{\alpha}\bar{\beta} & 0 & (\bar{\alpha}\bar{\beta})^2 & \dots \end{pmatrix} \\
\hat{\rho} \sim \mathcal{N}' = 3 &\rightarrow \begin{pmatrix} \alpha\beta & 0 & \beta\bar{\beta} & 0 & \bar{\alpha}\bar{\beta} & 0 & \dots \end{pmatrix} \\
\mathcal{N}' = 4 &\rightarrow \begin{pmatrix} 0 & \alpha\beta & 0 & \beta\bar{\beta} & 0 & \bar{\alpha}\bar{\beta} & \dots \end{pmatrix} \\
&\vdots \quad \begin{pmatrix} \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \\
&\quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\
\mathcal{N} = 1, 2, 3, 4, 5, 6, \dots &
\end{aligned} \tag{5.8}$$

Hence, a typical density matrix element between states with a difference in the total occupation number  $\Delta\mathcal{N}$  is suppressed by the product of  $\frac{\Delta\mathcal{N}}{2}$  factors of  $\mu_*^{-2} j_*^{-1}$ .

As already stated, the off-diagonal contribution to  $\rho_H$ —the *single-particle* density matrix—is suppressed by a factor  $\mu_*^{-2}(j - j')^{-1} \sim (y_*\varepsilon/T)^{-2}(j - j')^{-1}$ . Since the final result, when traced over all the frequency modes, should not depend on the parameter  $\varepsilon$ , we can choose it to be of the order of the BH temperature,  $\varepsilon \sim T$ . This amounts to saying that all the modes have the same frequency and, hence,  $j \sim \mathcal{O}(1)$ . It is then clear that the off-diagonal contribution to  $\rho_H$  is suppressed by at least  $y_*^{-2} \sim S_{BH}^{-2}(0) < N^{-2}$  with respect to the diagonal elements and is, therefore, inconsequential to the moments of  $\rho_H$ .

Nevertheless, the question of whether the off-diagonal elements can make a substantial correction to the moments of the *full* Hawking density matrix  $\hat{\rho}_H$  is well founded, as the number of off-diagonal elements is certainly very large. To address this concern, let us consider the regime in which  $N \sim S_{BH}$  because this is when significant changes to the nature of the BH radiation could be expected. To estimate the importance of the corrections for the Rényi entropy  $\frac{\text{tr}\hat{\rho}_H^2}{(\text{tr}\hat{\rho}_H)^2}$  we shall replace the hierarchical structure of the density matrix (5.8) by a uniform estimate of the typical value of the off-diagonal corrections. We do so by setting  $|\Delta\mathcal{N}| \sim N \sim S_{BH}$  for all the off-diagonal elements, so that the relative suppression factor goes as  $(\mu_*^2 j_*)^{-\frac{N}{2}}$ . The overall off-diagonal contribution to the Rényi entropy can now be calculated as the product of the dimensionality of the matrix  $d$  and the square of the relative suppression factor,  $d \times (\mu_*^2 j_*)^{-N}$ .

The size  $d$  of the full density matrix for  $N$  emitted particles is given by

$$d = \sum_{n=1}^N d_n = \sum_{n=1}^N \binom{n+W-1}{n} = \binom{N+W}{N}, \quad (5.9)$$

where  $d_n$  is the size of the  $n$ -particle subspace defined in Eq. (5.7), and  $W$  is the number of available frequencies. Since  $W \sim N$  for BH radiation (see Subsection 4.3), this gives  $d \sim e^{2N}$ .

Thus, the overall off-diagonal contribution, relative to that of the diagonal, is  $e^{2N}(y_*^2)^{-N} \sim e^{2N}N^{-2N} \ll e^{-N} \sim e^{-S_{BH}}$ , where we have again used that  $\varepsilon \sim T$  and  $j_* \sim \mathcal{O}(1)$ , leading to the estimate  $\mu_*^2 j_* \sim y_*^2 \sim S_{BH}^2 \sim N^2$ . It can be concluded that the off-diagonal contribution to the Hawking density matrix is exponentially suppressed. This conclusion agrees with the analysis of Mathur [25], who showed that small corrections at the level of the two-point function cannot change any of Hawking's basic outcomes. The physical reason behind this result is that the thermal radiation at future null infinity originates from ultra-high frequency modes at past null infinity which are highly red-shifted due to the presence of the horizon. We can therefore deduce that the existence of a region of high redshift—the horizon—induces a characteristic exponential suppression of the off-diagonal elements in mode-occupation number of the full density matrix.

Finally, let us use Eqs. (3.9) and (4.25), to determine the Rényi entropy for the Hawking model. As off-diagonal corrections are suppressed at least by order  $e^{-S_{BH}}$  at all times and the diagonal is approximately uniform, the single-particle density matrix always has  $N$  non-zero eigenvalues, each of which is unity up to negligible corrections. Hence, the estimate

$$(H_2)_{\text{Hawking}} \simeq N \ln 3 \quad (5.10)$$

is valid at all times.

## 5.2 The Page model

In the Page model of BH evaporation [10], the BH and the emitted radiation are assumed to be in a pure state in some large Hilbert space  $\mathcal{H}$ . This space is partitioned into an “in” part,  $\mathcal{H}_{BH}$ , representing the BH (including the ingoing radiation) and an “out” part,  $\mathcal{H}_{\text{out}}$ , representing the outgoing Hawking radiation, so that  $\mathcal{H} = \mathcal{H}_{BH} \otimes \mathcal{H}_{\text{out}}$ . The two Hilbert spaces are characterized only by their dimensionality.

Let us label the states in  $\mathcal{H}_{BH}$  by  $i = 1, \dots, m$ , where  $m = e^{S_{BH}}$ , and the states in  $\mathcal{H}_{\text{out}}$  by  $A = 1, \dots, n$ , where  $n = e^{S_{rad}}$ . A state in  $\mathcal{H}$  can then be written as  $|A, i\rangle$ . The working assumption is that both Hilbert spaces are large,  $n, m \gg 1$ . Therefore, the Bekenstein–Hawking entropy  $S_{BH}$  of the BH and the radiation entropy  $S_{rad} \sim N$  (which is calculated as if the radiation was in a thermal state) characterize the dimensionalities of the corresponding Hilbert spaces rather than the real entropies

of the subsystems. Since the total state is pure, the actual entropy of the radiation and that of the BH are entirely due to entanglement and, thus, equal to one another at all times.

The partition between “in” and “out” is meant to mimic the horizon separating the interior of the BH from its outside. The variation in time of the dimensions of the Hilbert spaces  $\mathcal{H}_{\text{BH}}$  and  $\mathcal{H}_{\text{out}}$  is meant to model the evaporation of the BH, so that  $\mathcal{H}_{\text{BH}}$  shrinks and  $\mathcal{H}_{\text{out}}$  grows while the number of emitted particles increases. However, additional physical effects resulting from the existence of a horizon are not taken into account.

The density matrix of the Page model  $\hat{\rho}_P$  is given by

$$\hat{\rho}_P = \rho_{A,i;B,j} |A, i\rangle \langle B, j| . \quad (5.11)$$

The reduced density matrix for the radiation is then obtained by tracing over the BH Hilbert space, just as in Eq. (2.7),

$$\hat{\rho}_{\text{out}} = \text{tr}_{BH} \hat{\rho}_P . \quad (5.12)$$

Because the combined state is pure, there must be a special basis in which  $\hat{\rho}_P$  has only a single eigenvalue. This can be chosen, without loss of generality, to be given by  $\rho_{1,1;1,1}$ . It follows that the matrix elements  $\rho_{A,i;B,j}$  can be expressed in terms of a basis transformation matrix  $U$  acting on the original state vector  $|1, 1\rangle$ . Denoting the resulting vector as  $V_{A,i} = U_{A,i;1,1}$ , we have

$$\rho_{A,i;B,j} = U_{A,i;A',i'} \rho_{A',i';B',j'} U_{B',i';B,i}^\dagger = V_{A,i} V_{B,i}^* . \quad (5.13)$$

and

$$\langle A | \rho_{\text{out}} | B \rangle = \sum_i V_{A,i} V_{B,i}^* . \quad (5.14)$$

The vectors  $V_{A,i}$  are meant to be treated statistically and assumed to be random vectors of unit size with a uniform distribution on an  $mn$ -dimensional sphere.

The initial investigation of such a random system was conducted by Lubkin [11] and then improved by many subsequent investigations [26–28]. The final result is that the distribution of eigenvalues of  $\hat{\rho}_{\text{out}}$  was found to obey the Marchenko–Pastur (MP) law [29], which is the eigenvalue distribution for a certain ensemble of semi-positive-definite, random matrices.

The MP distribution is generically divided into two parts, one set of large eigenvalues and another set of vanishingly small eigenvalues. It depends on only two parameters, the total number of the eigenvalues and the participation ratio corresponding to the number of large eigenvalues. The distribution of the eigenvalues (up to a normalization convention) is given by

$$\begin{aligned} P_{MP}(\lambda) d\lambda &= \left( \max(1 - c, 0) \delta(\lambda) + \frac{c}{2\pi\lambda} \sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)} \right) d\lambda \\ &= \left( \max(1 - c, 0) \delta(\lambda) + \frac{c}{2\pi\lambda} \sqrt{4/c - (\lambda - (1 + 1/c))^2} \right) d\lambda , \end{aligned} \quad (5.15)$$

where

$$\lambda_+ = (1 + 1/\sqrt{c})^2, \quad \lambda_- = (1 - 1/\sqrt{c})^2, \quad (5.16)$$

and  $c > 0$  is a parameter that determines the number of large eigenvalues.

When  $c \gg 1$ , all the eigenvalues are about equal and there are no zero eigenvalues whereas, for  $c < 1$ , a fraction  $1 - c$  of the eigenvalues is vanishing. For a matrix of dimension  $n$  with its eigenvalues distributed according to the MP law, the participation ratio PR, as defined in Eq. (4.27), is given by

$$PR = \frac{nc}{1+c}. \quad (5.17)$$

The higher moments of the MP distribution are also known in terms of  $n$  and  $c$  (see Appendix D).

For the Page model, in particular,

$$c = \frac{m}{n}, \quad (5.18)$$

and so

$$PR = \frac{mn}{m+n}. \quad (5.19)$$

The early times of BH evaporation, when  $n \ll m$ , correspond to  $c \gg 1$  and  $PR \sim n$ . Essentially, all the eigenvalues are about equal and their value is determined by the normalization convention. This means that the Rényi entropy of the radiation is the same as that of the Hawking model,  $(H_2)_{\text{Page}} \sim N \ln 3$ . But well after the Page time, when  $n \gg m$ , one finds that  $c \ll 1$  and  $PR \sim m \ll n$ . In this case, about  $n - m \sim n$  of the eigenvalues become zero and both reduced density matrices for the radiation and BH are becoming pure.

The Rényi entropy of the radiation for the Page model coincides with that of BH and is given by

$$H_2 = \ln PR \simeq \begin{cases} \ln n = S_{\text{rad}} \sim N & n < m, \\ \ln m = S_{\text{BH}}(N) = S_{\text{BH}}(0) - N & n > m. \end{cases} \quad (5.20)$$

This is depicted in Figure 1 as a function of the number of emitted particles. Since the entropy of the subsystems is entirely due to entanglement, it is symmetric if one exchanges the Hilbert spaces of the BH and radiation. As a consequence the evolution of the entropy is symmetric under reflections around the Page time, when the sizes of  $\mathcal{H}_{\text{BH}}$  and  $\mathcal{H}_{\text{out}}$  are equal; that is, when half of the particles have been emitted,  $N_{\text{Page}} \approx S_{\text{BH}}(0)/2$ .

### 5.3 The semiclassical model

The semiclassical model [13, 14] improves upon Hawking's framework by taking into account the BH's quantum fluctuations, as well as its time dependence due to the

emission of the radiation. In this setup, the single-particle density matrix  $\rho_{SC}$  is no longer diagonal and, as a result, the evaporation process becomes unitary even though the thermal-like emission spectrum is kept.

The basic prescription is to assign a quantum wavefunction to the collapsing shell of matter in Hawking's model and then recalculate all relevant quantities as expectation values. The main outcome is that  $\rho_{SC}$  picks up off-diagonal contributions that are uniform in terms of frequency but suppressed relative to the diagonal elements by  $C_{BH}^{1/2}(N)$ , where  $C_{BH}(N) = S_{BH}^{-1}(N) \ll 1$  is a classicality parameter—a “time-dependent  $\hbar$ ” that keeps track of how close the system is to the limit of a classical spacetime.

The elements of  $\rho_{SC}$  do have a non-uniform suppression in terms of emission time; modes emitted at different times tend to decohere. Nonetheless, if the radiation is being regularly monitored at intervals of  $\Delta N \sim \sqrt{S_{BH}}$  or less, then this suppression can be compensated. We will specifically be considering this case, which has been called the “tracking case” in [17]. Hence, the off-diagonal elements of the single-particle density matrix  $\rho_{SC}$  can be regarded as uniform in magnitude with respect to both frequency and emission time.

As explained in Subsection 4.3,  $\rho$  can be viewed as an  $N \times N$  matrix, with the indices running over the wave-packet modes with non-vanishing occupation number and with the diagonal elements given by the average occupation number for each mode. The elements of the semiclassical single-particle density matrix are found to be [13, 14]

$$\begin{aligned} (\rho_{SC})_{ii} &= 1 , \\ (\rho_{SC})_{i \neq j} &= \sqrt{C_{BH}(N)} e^{i\theta_{ij}} , \end{aligned} \tag{5.21}$$

where the phases  $\theta_{ij}$  can be treated as random for most purposes (but see below).

Let us next consider the full density matrix  $\hat{\rho}_{SC}$  for this model. In [13] (see, in particular, Sect. 2.4), it was argued that the off-diagonal elements in mode-occupation number of the density matrix will be suppressed much in the same way as for the Hawking model, as discussed in Subsection 5.1. The physical reason for this is the same as before—the presence of the horizon implies that the frequencies observed at future null infinity are highly red-shifted and, thus, determine the magnitude of the suppression factor of the off-diagonal elements. The technical explanation is as follows: In the semiclassical model, the terms that lead to off-diagonal elements in mode-occupation number are, again, terms of the form  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  as in Eq. (5.2). These contain a classical term, which is equal to the Hawking-model contribution, along with a semiclassical correction. Irrespective of the details, it is clear that the semiclassical corrections to  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  vanish as some power of  $C_{BH}$  in the limit  $C_{BH} \rightarrow 0$ . This is enough to guarantee the exponential suppression of the semiclassical contribution to the off-diagonal elements in mode-occupation number.

Hence, there will be a hierarchical structure in the suppression factors of the off-diagonal elements in mode-occupation number, as shown in (5.8). Like before, the off-diagonal elements can be expressed in terms of  $(\mathcal{N} + \mathcal{N}')$ -point functions, which will factorize into a product of  $\frac{\mathcal{N} + \mathcal{N}'}{2}$  two-point functions. The strength of an  $(\mathcal{N} + \mathcal{N}')$ -point function will then be suppressed by a factor similar to the suppression factor of the Hawking model.

In some sense, this semiclassical model can be viewed as the “middle ground” between the other two models; on one hand, remaining almost thermal like Hawking’s with the associated hierarchical structure of the density matrix but, on the other, evolving over time like Page’s. At the early stages of BH evaporation, the semiclassical model is essentially Hawking’s plus small corrections. However, at later times (after the Page time), the dominant contributions will come from those elements that are off-diagonal in frequency. This can be attributed to the effective perturbative parameter being  $NC_{\text{BH}}$  for this framework [13], as this parameter grows monotonically throughout the evaporation process and finally becomes large ( $NC_{\text{BH}} > 1$ ) at times later than the Page time. It is the large size of this effective perturbation parameter that allows the semiclassical model to evade the conclusions of Mathur [25]; in particular, it leads to a Rényi entropy for the radiation that is decreasing in time (see below).

### 5.3.1 Semiclassical Rényi entropy

The objective here is to estimate the Rényi entropy of the BH radiation  $H_2(N)$  using the single-particle density matrix  $\rho_{SC}$  as specified in Eq. (5.21). However, we are currently lacking a precise quantitative knowledge of  $\rho_{SC}$ . In particular, in order to use Eq. (3.9) for  $H_2(N)$ , one needs to know the eigenvalue distribution  $P(\lambda)$  of  $\rho_{SC}$ :

$$H_2(N) = \langle \text{tr} [\ln (1 + 2\rho)] \rangle_{P(\lambda)} = N \int_0^\infty \ln (1 + 2\lambda) P(\lambda) d\lambda. \quad (5.22)$$

This deficiency can be attributed to our lack of specification as to the exact nature of the phases  $e^{i\theta_{ij}}$ . These phases are not entirely random but rather dictated by the effect of the quantum horizon on correlations in the emitted radiation [14]. In this subsection, we will instead call upon the known properties of  $\rho_{SC}$  in order to draw conclusions about the entropy of the radiation in the semiclassical model.

We expect that  $\rho_{SC}$  has some number of large eigenvalues with the rest vanishing, whereby the large eigenvalues are distributed about the mean in a way that depends on their total number. This is most clearly seen from the participation ratio  $PR$  of the single-particle density matrix  $\rho_{SC}$ . The participation ratio provides, for a given matrix, a measure of the number of non-vanishing eigenvalues (e.g., [23]). For  $\rho_{SC}$ , it is given by [17]

$$PR = \frac{(\text{tr} \rho_{SC})^2}{\text{tr} \rho_{SC}^2} = \frac{N}{1 + NC_{\text{BH}}} = \frac{N(S_{BH}(0) - N)}{S_{BH}(0)}. \quad (5.23)$$

Then, since  $\text{tr } \rho_{SC} = N$  (as shown in Subsection 4.3), the average value of a non-vanishing eigenvalue can be estimated as

$$\bar{\lambda} = \frac{\text{tr } \rho_{SC}}{PR} = \frac{N}{PR} = 1 + NC_{\text{BH}}. \quad (5.24)$$

This finding reveals that, up to the Page time ( $NC_{\text{BH}} < 1$ ), there are about  $N$  eigenvalues of order one. The properties of the emitted radiation in this regime coincide with the predictions of the Hawking model: Each of the  $N$  available modes has an occupation number of order one and the radiation entropy grows linearly with the number of emitted particles.

On the other hand, after the Page time ( $NC_{\text{BH}} > 1$ ), the participation ratio (5.23) stops growing and eventually becomes much smaller than  $N$ , so that the average  $\bar{\lambda}$  is much larger than one. This signals that the matrix  $\rho_{SC}$  has a small fraction of large eigenvalues, with the majority of eigenvalues vanishing. Physically, it means that, out of the  $N$  available radiation modes, there is only a small fraction with large occupation numbers. In this regime, the entropy of the radiation (5.22) is dominated by the contribution from the occupied modes and is smaller than the entropy of the thermal radiation (5.10) which is predicted by the original Hawking model:

$$H_2(N) \simeq PR \ln(1 + 2\bar{\lambda}) \sim PR \ln\left(\frac{N}{PR}\right) \ll N, \quad \text{for } PR \ll N. \quad (5.25)$$

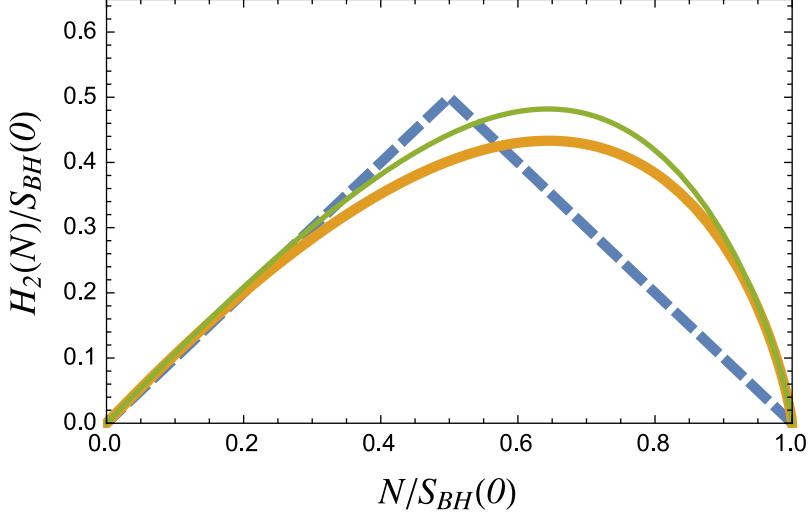
Therefore, the known properties of the single-particle density matrix  $\rho_{SC}$  guarantee that the radiation entropy in the semiclassical model does not grow steadily with number of emitted particles, as predicted by the Hawking computation, but rather starts to decrease after the Page time.

In order to illustrate the evolution of the radiation entropy in the semiclassical model, let us consider two examples for the eigenvalue distribution of  $\rho_{SC}$ , both having the property that the participation ratio can be much smaller than the dimensionality of the matrix. First, we consider the simplest case when  $\rho_{SC}$  has exactly  $PR$  non-vanishing eigenvalues of magnitude  $\bar{\lambda}$  and the remaining  $(N - PR)$  eigenvalues are exactly vanishing. The corresponding distribution function reads

$$P(\lambda) = \frac{N - PR}{N} \delta(\lambda) + \frac{PR}{N} \delta\left(\lambda - \frac{N}{PR}\right), \quad (5.26)$$

where the participation ratio is that of Eq. (5.23). The Rényi entropy in such a case is given by

$$\begin{aligned} H_2(N) &= PR \ln(1 + 2\bar{\lambda}) = \frac{N}{1 + NC_{\text{BH}}} \ln(3 + 2NC_{\text{BH}}) \\ &= S_{BH}(0) \frac{N}{S_{BH}(0)} \left(1 - \frac{N}{S_{BH}(0)}\right) \ln\left(1 + \frac{2}{1 - N/S_{BH}(0)}\right). \end{aligned} \quad (5.27)$$



**Figure 1.** Dependence of the Rényi entropy  $H_2$  of the BH radiation on the number of emitted particles  $N$  for the semiclassical model in Eq.(5.21) for two eigenvalue distributions. First, assuming that the non-vanishing eigenvalues of the two-point function are distributed according to Eq. (5.26) (solid, thick orange) and then to the Marchenko–Pastur law in Eq. (5.15) (solid, thin green). For comparison, the dashed blue line depicts the prediction of the Page model from Eq. (5.20).

The evolution of this radiation entropy with the number of emitted particles is presented in Figure 1, which shows the predicted decline after the Page time.

Another eigenvalue distribution that illustrates the properties of  $\rho_{SC}$  is the MP distribution (5.15), as already discussed in the context of the Page model. The crucial difference from the Page model itself is that we now use the MP distribution to model the single-particle density matrix  $\rho_{SC}$  and not the reduced density matrix which represents the state of the emitted radiation. Requiring the participation ratio of the MP-distributed matrix (5.17) to be equal to the semiclassical prediction (5.23), one finds that the parameter  $c$  of the MP distribution is given by

$$c = \frac{1}{NC_{\text{BH}}} . \quad (5.28)$$

The Rényi entropy of the radiated particles can then be obtained by plugging the MP distribution (5.15) into Eq. (5.22). The evolution of the entropy is shown in Figure 1. One can observe that the entropy is decreasing after the Page time, just as expected for the semiclassical model.

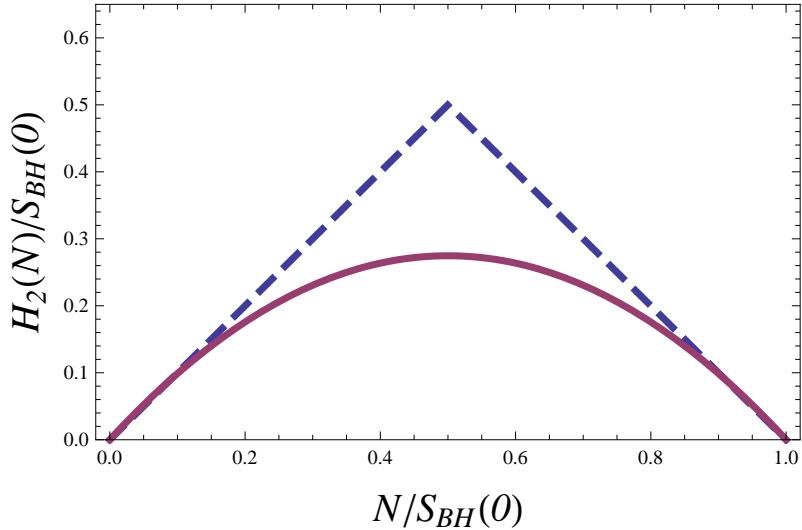
Although these two examples were used for illustrative purposes, we do expect the eigenvalue distribution for  $\rho_{SC}$  to be similar in some aspects to the MP distribution. The moments of the distribution function of the eigenvalues of  $\rho_{SC}$  are calculated and compared to the moments of the MP distribution in Appendix D. In spite of differing in details, the two sets of moments follow a similar pattern.

Having established that the Rényi entropy of the radiation in the semiclassical model starts to decrease after the Page time, let us speculate on a viable form for its functional dependence on the number of emitted particles. What is known about the semiclassical model is that  $H_2(\rho_{SC})$  (1) must tend to the Hawking value of  $N \ln 3$ , as in Eq. (5.10), at early times and (2) must be symmetric under an exchange between the radiation and the BH, since the two reduced density matrices necessarily share the same set of eigenvalues. In particular, all physical quantities are required to be symmetric under the exchange of  $N \leftrightarrow S_{BH}(N)$ , as long as these quantities characterize the sizes of their respective Hilbert spaces  $\mathcal{H}_{\text{out}}$  and  $\mathcal{H}_{\text{BH}}$ .

The participation ratio (5.23) is manifestly symmetric in just this way,

$$PR(\rho_{SC}) = \frac{N}{1 + NC_{\text{BH}}(N)} = \frac{NS_{BH}(N)}{S_{BH}(N) + N}. \quad (5.29)$$

It is also of the same form as the participation ratio of the MP distribution (5.17), given that one identifies the parameters  $n$  and  $m$  of the Page model to be  $N$  and  $S_{BH}(N)$ , respectively.<sup>9</sup> This may seem peculiar because, as discussed in Appendix D, the actual eigenvalue distribution of  $\rho_{SC}$  differs from the MP law. But, as far as we can tell,  $\frac{NS_{BH}(N)}{S_{BH}(N) + N}$  is the simplest non-trivial, symmetric function which tends to  $N$  for  $N \ll S_{BH}(0)$ , suggesting that such an identification is valid.



**Figure 2.** Dependence of the Rényi entropy  $H_2$  of the BH radiation on the number of emitted particles  $N$ . The solid purple line depicts the prediction of the semiclassical model from Eq. (5.30) and the dashed blue line depicts the prediction of the Page model from Eq. (5.20).

<sup>9</sup>Note that, in the Page model,  $n$  and  $m$  are the dimensions of the reduced density matrices  $\hat{\rho}_{\text{out}}$  and  $\hat{\rho}_P$ , and are exponentially large numbers,  $n \sim e^N$  and  $m \sim e^{S_{BH}(N)}$ .

Motivated by these observations, our expectation is that the semiclassical Rényi entropy can be expressed in terms of the participation ratio as follows:

$$H_2(\rho_{SC}) = PR(\rho_{SC}) \left[ \ln 3 + c_{-1} \frac{1}{PR(\rho_{SC})} + c_{-2} \frac{1}{(PR(\rho_{SC}))^2} + \dots \right], \quad (5.30)$$

where the coefficients are dimensionless numbers that are determined by the higher moments of the eigenvalue distribution. To see that  $c_0 = \ln 3$  is indeed the correct leading-order coefficient, consider that, at early times,  $N \simeq PR$  to leading order in a  $1/PR$  expansion. Thus, the average eigenvalue  $\bar{\lambda} = \frac{N}{PR} \simeq 1$  and the logarithm in  $H_2$  then expands as  $\ln(1 + 2\bar{\lambda}) \simeq \ln 3$ .

The dependence (5.30) of the Rényi entropy on the number of emitted particles for a semiclassical BH is presented in Figure 2. Because the Rényi entropy depends only on the participation ratio  $\frac{N}{1+NC_{BH}}$ , one can determine the evolution of the former from that of the latter. The participation ratio steadily grows until reaching a maximum at the Page time and then steadily declines for the remainder of the evaporation process. This is qualitatively similar behavior to the Rényi entropy for the Page model although, in the semiclassical model, the rate of purification—which is the rate of deviation from the linearly growing result of the Hawking's model—is actually faster than the Page-model rate [17].

## 6 Conclusion

We have exposed the relationship between the density matrix of the outgoing radiation from a BH and the corresponding single-particle density matrix in the case that the density matrix is approximately diagonal in mode-occupation number. It was then shown that the presence of the horizon leads to a high suppression of the off-diagonal elements in mode-occupation number of the density matrix for the emitted radiation. We have therefore concluded that the density matrix is approximately diagonal in mode-occupation number at all times. We have also shown how to regularize the infinities which arise as a consequence of Hawking's idealized picture of an eternal BH. It was explained how the state of the emitted radiation from a BH is different from that of a standard black body of the same temperature. This analysis was then applied to three models of BH evaporation as a means of clarifying their differences and contrasting their main features. Let us briefly summarize the main observations:

### The Hawking model:

The off-diagonal elements of the single-particle density matrix are highly suppressed for different frequencies and emission times, and those of the full density matrix are exponentially suppressed for differences in occupation numbers. The

Rényi entropy (4.20) scales with the number of emitted particles  $N$  at all times, and so there is no possibility for the radiation to be purified.

### The Page Model:

For a randomly chosen basis whose relationship to Hawking's basis is unspecified, the off-diagonal elements of the density matrix are uniform in magnitude with random phases. There is no hierarchy between different elements of the matrix. The Rényi entropy of the radiation (5.20) increases linearly as  $N$  until the Page time and decreases linearly as  $S_{BH}(0) - N$  afterwards. Hence, purification is inevitable.

### The semiclassical model:

The off-diagonal elements of the single-particle density matrix are suppressed by a power of  $S_{BH}^{-1/2}$  in a uniform way, whereas those of the density matrix are exponentially suppressed for different mode-occupation numbers in a similar way to the Hawking model. At the Page time, the contribution from the off-diagonal elements with regard to frequency and emission time becomes significant because of a perturbative parameter that grows monotonically with time. The Rényi entropy (5.30) scales with the participation ratio of the single-particle density matrix, which increases until the Page time and decreases thereafter. This suggests that the radiation starts to purify and does so at a rate which is faster than that of the Page model. The question of how unitarity is recovered for a process of gravitational collapse in the semiclassical model is discussed in detail in a companion paper [17].

## Acknowledgments

The research of LA, RB, AK was supported by the Israel Science Foundation grant no. 239/10. The work of AK was also supported by the Kreitman foundation. The work of LA was also supported by the Minerva Foundation. The research of AJMM received support from an NRF Incentive Funding Grant 85353, an NRF Competitive Programme Grant 93595 and Rhodes Research Discretionary Grants. AJMM thanks Ben-Gurion University for their hospitality during his visit.

## A Expression for the full density matrix in terms of $\rho_j^i$

In order to find an expression for the density matrix, it is useful to calculate a vacuum expectation value; namely,

$$\langle 0_- | e^{b_i^+ \mu^i} e^{\lambda_j b^j} | 0_- \rangle. \quad (\text{A.1})$$

The resulting function of  $\mu^i$  and  $\lambda_j$  can be used as a generating function for the expectation values of any normal-ordered powers of the creation and annihilation

operators  $b_i^+$  and  $b^j$ . The expectation value (A.1) can be calculated by using (3.2) and applying the Baker–Campbell–Hausdorff (BCH) formula for the exponents:

$$\langle 0_- | e^{b_i^+ \mu^i} e^{\lambda_j b^j} | 0_- \rangle = \langle 0_- | e^{\mu^i \alpha_i^j a_j^+} e^{-\mu^i \beta_{ij} a^j} e^{-\lambda_j \bar{\beta}^{jk} a_k^+} e^{\lambda_j \bar{\alpha}_k^j a^k} | 0_- \rangle e^{-\frac{1}{2} \mu^i \alpha_i^j \beta_{kj} \mu^k} e^{-\frac{1}{2} \lambda_i \bar{\alpha}_k^i \bar{\beta}^{jk} \lambda_j}. \quad (\text{A.2})$$

The exponential factors on the outside of the average arise from the commutators. By using Hawking's expressions in [2] for the Bogolyubov coefficients  $\alpha_{ij}$  and  $\beta_{ij}$ , one can show that

$$\alpha_i^j \beta_{kj} = 0, \quad \bar{\alpha}_k^i \bar{\beta}^{jk} = 0, \quad (\text{A.3})$$

and so the external exponentials are equal to unity as their exponents are vanishing.

Relations (A.3) also ensure that only the matrix elements of  $\hat{\rho}^{\text{out}}$  between states with the same total occupation number are non-vanishing. But, as discussed in Section 5, these relations hold only approximately for realistic BHs.

Returning to Eq. (A.2), one can see that the two outer factors within the expectation value disappear after acting on the vacuum  $|0_-\rangle$ . The remaining expression can be further simplified by applying the BCH relation once again:

$$\langle 0_- | e^{-\mu^i \beta_{ij} a^j} e^{-\lambda_j \bar{\beta}^{jk} a_k^+} | 0_- \rangle = e^{\lambda_i \bar{\beta}^{ik} \beta_{jk} \mu^j} \cdot \langle 0_- | e^{-\lambda_j \bar{\beta}^{jk} a_k^+} e^{-\mu^i \beta_{ij} a^j} | 0_- \rangle = e^{\lambda_i \bar{\beta}^{ik} \beta_{jk} \mu^j}. \quad (\text{A.4})$$

Finally, the answer for the generating function (A.1) reads

$$\langle 0_- | e^{b_i^+ \mu^i} e^{\lambda_j b^j} | 0_- \rangle = e^{\lambda_i \rho_j^i \mu^j}, \quad (\text{A.5})$$

where  $\rho_j^i$  is the single-particle density matrix as defined in Eq. (3.3). For Hawking,  $\rho_j^i$  is a diagonal matrix but, in more general setups, it can also have non-zero off-diagonal elements. In any case,  $\rho_j^i$  is Hermitian and can thus be diagonalized by a unitary transformation.

We next look for a closed expression for the density matrix  $\hat{\rho}^{\text{out}}$  in terms of the  $b$ 's. In analogy with the harmonic oscillator, let us try the *ansatz*

$$\hat{\rho}^{\text{out}} = \frac{1}{Z} e^{-b_i^+ \Omega^i_j b^j}, \quad (\text{A.6})$$

where  $\Omega$  is some  $c$ -number Hermitian matrix and  $Z$  is the normalization factor. The matrix  $\Omega$  can be diagonalized by a unitary transformation,  $\Omega = U^+ D U$ , with  $D$  being a real diagonal matrix. This allows us to introduce new annihilation operators  $d^j \equiv U^j_i b^i$  for which the quadratic form  $b^+ \Omega b = b^+ U^+ D U b = d^+ D d$  is diagonal. Such a transformation is canonical, so that  $\{d_i^+, d_i\}$  are also a good set of creation and annihilation operators.

Moreover, since this transformation does not mix the creation and annihilation operators between each other, the subspaces with a fixed total occupancy of  $b$ - and  $d$ -particles coincide, and the traces over the Fock spaces of  $b$ - and  $d$ -particles are

equal. In particular, one can find the normalization  $Z$  by summing over the states with a fixed occupation number of  $d$ -particles  $|n_i\rangle$ :

$$\begin{aligned} Z = \text{tr} \left[ e^{-b_i^+ \Omega^i_j b^j} \right] &= \text{tr} \left[ e^{-\sum_i d_i^+ D_i^i d^i} \right] = \prod_i \sum_{n_i=0}^{\infty} \langle n_i | e^{-\sum_i D_i^i d_i^+ d^i} | n_i \rangle = \\ &= \prod_i \frac{1}{1 - e^{-D_i^i}} = \det \left[ \frac{1}{1 - e^{-D}} \right] = \det \left[ \frac{1}{1 - e^{-\Omega}} \right]. \quad (\text{A.7}) \end{aligned}$$

The last equality is due to the invariance of the determinant under unitary transformations of the matrix.

According to *ansatz* (3.4), the corresponding generating function (A.1) is

$$\langle e^{b^+ \mu} e^{\lambda b} \rangle_{\hat{\rho}} \equiv \frac{1}{Z} \text{tr} \left[ e^{b^+ \mu} e^{\lambda b} e^{-b^+ \Omega b} \right]. \quad (\text{A.8})$$

If it is possible to choose  $\Omega$  so that the result matches (A.5), then *ansatz* (A.6) gives the correct expression for the density matrix.

As before, the generating function (A.8) can be calculated by tracing in the  $d$ -particle basis:

$$\begin{aligned} \frac{1}{Z} \text{tr} \left[ e^{b^+ \mu} e^{\lambda b} e^{-b^+ \Omega b} \right] &= \frac{1}{Z} \text{tr} \left[ e^{d^+ U \mu} e^{\lambda U^+ d} e^{-d^+ D d} \right] = \\ &= \prod_i \frac{1}{Z_i} \sum_{n_i=0}^{\infty} \langle n_i | e^{d_i^+ (U \mu)^i} e^{(\lambda U^+)_i d^i} e^{-d_i^+ D_i^i d^i} | n_i \rangle, \quad (\text{A.9}) \end{aligned}$$

where each of the factors in the bottom line corresponds to a single oscillator.

In order to evaluate the above expression, we will use the identity

$$(1 - e^{-D}) \text{tr} \left[ e^{a^+ \mu} e^{\lambda a} e^{-D a^+ a} \right] = \exp \left( \frac{\lambda \mu}{e^D - 1} \right), \quad (\text{A.10})$$

which is derived in [30]. (For a web resource, see, *e.g.*, [31].) The expectation value (A.8) then reads

$$\begin{aligned} \langle e^{b^+ \mu} e^{\lambda b} \rangle_{\hat{\rho}} &= \prod_i \exp \left( \frac{(\lambda U^+)_i (U \mu)^i}{e^{D_i^i} - 1} \right) = \exp \left( \sum_i \frac{(\lambda U^+)_i (U \mu)^i}{e^{D_i^i} - 1} \right) \equiv \\ &\equiv \exp \left( \lambda U^+ \frac{1}{e^D - 1} U \mu \right) = \exp \left( \lambda \frac{1}{e^\Omega - 1} \mu \right). \quad (\text{A.11}) \end{aligned}$$

Comparing the last expression to (A.5), we find that the right-hand sides coincide if the matrix  $\Omega$  is related to  $\rho$  as

$$\rho = \frac{1}{e^\Omega - 1}, \quad \text{or} \quad e^\Omega = 1 + \frac{1}{\rho}. \quad (\text{A.12})$$

## B Grey-body factors

From the properties of the grey-body factors, one can see that the state of the field outside a large BH is quite different from the usual thermal state in an empty box,  $\hat{\rho} = \frac{1}{Z} e^{-\beta \hat{H}}$ . The actual state of the radiation is that of a field in thermal equilibrium with the BH, which is very selective in what it absorbs and emits. For example, the occupation number of the modes with high angular momentum will be much less than what one would naively expect from just the Boltzmann suppression.

One way to present this modification away from thermality is to notice that, by definition, the grey-body factor  $\Gamma_{\omega lm}$  is equal to the modulus square of the scattering amplitude of the corresponding mode on the BH. Therefore, it is physically intuitive to introduce a corresponding effective scattering cross-section (see, e.g., Ref. [22]),

$$\sigma_{\omega lm} = \frac{\pi}{\omega^2} \Gamma_{\omega lm} . \quad (\text{B.1})$$

The rate of emission can then be written as

$$\frac{N_{\omega lm}}{\delta t} = \int_0^\infty \frac{\omega^2 d\omega}{2\pi^2} \frac{\sigma_{\omega lm}}{e^{\omega/T} - 1} . \quad (\text{B.2})$$

This is a three-dimensional thermal emission rate, where the geometric area of the emitter is different for each mode and is given by the absorption cross-section. In this language, the suppression of the high- $l$  modes comes from the fact that the effective area of the BH as seen by them is much smaller than its geometric area.

Another way to see the difference from the usual thermal state is to notice that the density matrix of the equilibrium state has the form  $\hat{\rho} = Z^{-1} e^{-b^+ \Omega b}$  with  $\Omega = \omega/T$  only if  $\Gamma_{\omega lm} = 1$ . Indeed, the relation (3.5) between the single-particle density matrix  $\rho = \frac{\Gamma_{\omega lm}}{e^{\omega/T} - 1}$  and the matrix  $\Omega$  implies that

$$\Omega_{\omega lm} = \ln \left( \frac{e^{\omega/T} - (1 - \Gamma_{\omega lm})}{\Gamma_{\omega lm}} \right) . \quad (\text{B.3})$$

Therefore, the density matrix for the radiation takes the form [cf. Eq. (2.10)]

$$\hat{\rho} = \frac{1}{Z} \prod_{\omega lm} e^{-\Omega_{\omega lm} \hat{N}_{\omega lm}} = \frac{1}{Z} \prod_{\omega lm} \left( \frac{\Gamma_{\omega lm}}{e^{\omega/T} - (1 - \Gamma_{\omega lm})} \right)^{\hat{N}_{\omega lm}} . \quad (\text{B.4})$$

For each given mode, the ratio of the probability of having  $(N + 1)$  particles to the probability of having  $N$  particles is constant and given by

$$\frac{P(N + 1)}{P(N)} = \frac{\Gamma_{\omega lm}}{e^{\omega/T} - (1 - \Gamma_{\omega lm})} . \quad (\text{B.5})$$

Meaning that, apart from the overall Boltzmann suppression factor, the emitted particle has a probability of  $(1 - \Gamma)$  to be scattered back into the BH,

$$P(N + 1) = e^{-\omega/T} [\Gamma P(N) + (1 - \Gamma) P(N + 1)] . \quad (\text{B.6})$$

## C Off-diagonal corrections to the Hawking model

In order to estimate the off-diagonal corrections to the Hawking density matrix, we first need to consider the corrections to the products

$$\bar{\beta}^{ik} \beta_{jk} , \quad \alpha_i^j \beta_{kj} , \quad \bar{\alpha}_k^i \bar{\beta}^{jk} , \quad (C.1)$$

where the second index is summed over. The first product defines the standard single-particle density matrix  $\rho_j^i \equiv \bar{\beta}^{ik} \beta_{jk}$  and, in the case of Hawking's calculation, is diagonal and completely determines the density matrix of the outgoing radiation (3.4). The products  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  arise in the expectation values of operators like

$$\langle 0_- | (b^+)^n b^{n+2k} | 0_- \rangle \sim (\beta \bar{\beta})^n \cdot (\alpha\beta)^k , \quad (C.2)$$

$$\langle 0_- | (b^+)^{n+2k} b^n | 0_- \rangle \sim (\beta \bar{\beta})^n \cdot (\bar{\alpha}\bar{\beta})^k \quad (C.3)$$

and, thus, define the elements of the density matrix between states that differ in total occupation number,  $\Delta\mathcal{N} = 2k$ .

The products  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  also appear during the computation of the generating function (A.2). Consequently, the sole dependence of the density matrix on the single-particle density matrix  $\rho_j^i$  is only truly valid in the case when the products  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  vanish. This assumption is true for Hawking's idealized calculation but, as shown below, modified for physically realistic BHs.

In order to find the corrections to the products (C.1), we rewrite the definition of the coefficients  $\alpha_{\omega\omega'}$  and  $\beta_{\omega\omega'}$  in the basis of Fourier modes,

$$\alpha_{\omega\tilde{\omega}} = i f_\alpha(\omega) (\tilde{\omega})^{-1/2 + \frac{i\omega}{\kappa}} e^{i(\omega-\tilde{\omega})v_0} , \quad (C.4)$$

$$\beta_{\omega\tilde{\omega}} = -i f_\beta(\omega) (\tilde{\omega})^{-1/2 + \frac{i\omega}{\kappa}} e^{i(\omega+\tilde{\omega})v_0} , \quad (C.5)$$

where

$$f_\alpha(\omega) \equiv \frac{t_\omega}{2\pi} \frac{1}{\sqrt{\omega}} \Gamma\left(1 - \frac{i\omega}{\kappa}\right) e^{\frac{\pi\omega}{2\kappa}} , \quad f_\beta(\omega) \equiv f_\alpha(\omega) e^{-\frac{\pi\omega}{\kappa}} , \quad (C.6)$$

and  $v_0$  denotes the position of the BH horizon in advanced time,  $\kappa = 2\pi T$  is the BH surface gravity and  $t_\omega$  is the transmission coefficient for which  $|t_\omega|^2 = \Gamma_\omega$ .

The products (C.1) can be rewritten in the wave-packet basis of Section 4. For instance,

$$(\bar{\beta}\beta)^{jn}_{j'n'} = \varepsilon^{-1} \int_{j\varepsilon}^{(j+1)\varepsilon} d\omega e^{2\pi i n \omega / \varepsilon} \int_{j'\varepsilon}^{(j'+1)\varepsilon} d\omega' e^{-2\pi i n' \omega' / \varepsilon} \rho_{\bar{\beta}\beta}(\omega, \omega') , \quad (C.7)$$

where

$$\rho_{\bar{\beta}\beta}(\omega, \omega') \equiv \int_0^\infty d\tilde{\omega} \bar{\beta}_{\omega\tilde{\omega}} \beta_{\omega'\tilde{\omega}} . \quad (C.8)$$

There are similar expressions for  $(\alpha\beta)_{jn j'n'}$  and  $(\bar{\alpha}\bar{\beta})^{jn j'n'}$ . The product (C.8) and its analogues can be expressed in terms of the Fourier-mode basis as

$$\rho_{\bar{\beta}\beta}(\omega, \omega') = \bar{f}_\beta(\omega) f_\beta(\omega') e^{i(\omega' - \omega)v_0} \int_{-\infty}^{+\infty} dy e^{iy/\kappa(\omega' - \omega)}, \quad (\text{C.9})$$

$$\rho_{\alpha\beta}(\omega, \omega') = f_\alpha(\omega) f_\beta(\omega') e^{i(\omega' + \omega)v_0} \int_{-\infty}^{+\infty} dy e^{iy/\kappa(\omega' + \omega)}, \quad (\text{C.10})$$

$$\rho_{\bar{\alpha}\bar{\beta}}(\omega, \omega') = \bar{f}_\alpha(\omega) \bar{f}_\beta(\omega') e^{-i(\omega' + \omega)v_0} \int_{-\infty}^{+\infty} dy e^{-iy/\kappa(\omega' + \omega)}, \quad (\text{C.11})$$

where the integration variable has been changed to  $y = \ln(\tilde{\omega})$ .

Substituting the previous set of relations into the corresponding expressions for the products in the wave-packet basis and then integrating over the frequencies  $\omega$  and  $\omega'$ , we arrive at

$$(\bar{\beta}\beta)^{jn} = f_\beta(\bar{\omega}') \bar{f}_\beta(\bar{\omega}) \frac{2\kappa}{\pi} \int_{-\infty}^{+\infty} d\mu \frac{e^{2\pi i \mu(j' - j)}}{(\mu - \tilde{n})(\mu - \tilde{n}')} \sin^2(\pi\mu), \quad (\text{C.12})$$

$$(\alpha\beta)_{jn j'n'} = f_\alpha(\bar{\omega}') f_\beta(\bar{\omega}) \frac{2\kappa}{\pi} \int_{-\infty}^{+\infty} d\mu \frac{e^{2\pi i \mu(j' + j + 1)}}{(\mu - \tilde{n})(\mu - \tilde{n}')} \sin^2(\pi\mu), \quad (\text{C.13})$$

$$(\bar{\alpha}\bar{\beta})^{jn j'n'} = \bar{f}_\alpha(\bar{\omega}') \bar{f}_\beta(\bar{\omega}) \frac{2\kappa}{\pi} \int_{-\infty}^{+\infty} d\mu \frac{e^{-2\pi i \mu(j' + j + 1)}}{(\mu - \tilde{n})(\mu - \tilde{n}')} \sin^2(\pi\mu), \quad (\text{C.14})$$

where  $\bar{\omega} \equiv \varepsilon(j + 1/2)$  and  $\bar{\omega}' \equiv \varepsilon(j' + 1/2)$  are the mean values of the frequencies in the given range of integration,  $\tilde{n} \equiv n - n_0$ ,  $\tilde{n}' \equiv n' - n_0$  with  $n_0$  defined by  $v_0 = 2\pi n_0 \varepsilon^{-1}$ , and the integration variable has been changed to  $\mu = y\varepsilon/(2\pi\kappa)$ . We see that the three integrals differ only by the power in the exponent, and so it is convenient to define

$$J_{\bar{\beta}\beta} \equiv j' - j, \quad J_{\alpha\beta} \equiv j' + j + 1, \quad J_{\bar{\alpha}\bar{\beta}} \equiv -(j' + j + 1). \quad (\text{C.15})$$

Let us now introduce infrared and ultraviolet cutoffs for the integration range of the frequencies, as discussed in the Subsection 5.1,

$$\int_{-\infty}^{+\infty} d\mu \rightarrow \int_{-\mu_*}^{\mu_*} du. \quad (\text{C.16})$$

Due to the similarity of the three integrals above, we need only to consider one,

$$I_{\Delta n}(J) = \int_{-\mu_*}^{+\mu_*} d\mu \frac{e^{2\pi i \mu J}}{(\mu - \tilde{n})(\mu - \tilde{n}')} \sin^2(\pi\mu), \quad (\text{C.17})$$

and then interpret the results for the different choices of  $J$  in Eq. (C.15). The integral does need, however, to be treated differently for the cases with  $\Delta n \equiv n' - n = 0$  and  $\Delta n \neq 0$ , as well as for the cases with  $J = 0$  and  $J \neq 0$ .

### C.1 Off-diagonal elements in frequency

When  $\Delta n = 0$ , the integral (C.17) becomes

$$I_0(J) = \int_{-\mu_*}^{\mu_*} d\mu \frac{e^{2\pi i \mu J}}{(\mu - \tilde{n})^2} \sin^2(\pi \mu) . \quad (\text{C.18})$$

For the case  $J = 0$ , the previous integral can be evaluated to give

$$I_0(0) = -\frac{\sin^2(\pi \mu_*)}{\mu_* - \tilde{n}} - \frac{\sin^2(\pi \mu_*)}{\mu_* + \tilde{n}} + \pi \text{Si}(2\pi(\mu_* - \tilde{n})) + \pi \text{Si}(2\pi(\mu_* + \tilde{n})) , \quad (\text{C.19})$$

where  $\text{Si}(x) = \int_0^x dt \sin t/t$  is the sine integral. It has the large- $x$  expansion

$$\text{Si}(x) = \frac{\pi}{2} - \frac{\cos x}{x} - \frac{\sin x}{x^2} + \mathcal{O}(x^{-3}) . \quad (\text{C.20})$$

Hence, in the limit when  $\mu_* \gg \tilde{n}$ , the integral becomes

$$I_0(0) = \pi^2 - \frac{1}{\mu_*} - \frac{\sin(2\pi \mu_*)}{2\pi \mu_*^2} + \mathcal{O}(\mu_*^{-3}) . \quad (\text{C.21})$$

When  $J \neq 0$ , the integral can be split as

$$I_0(J) = \frac{1}{2} \left( h(J) - \frac{1}{2} h(J-1) - \frac{1}{2} h(J+1) \right) , \quad (\text{C.22})$$

where

$$h(J) = \int_{-\mu_*}^{\mu_*} d\mu \frac{e^{2\pi i \mu J}}{(\mu - \tilde{n})^2} . \quad (\text{C.23})$$

We find that the last integral is expressible as

$$h(J) = -\frac{e^{2\pi i \mu_* J}}{\mu_* - \tilde{n}} - \frac{e^{-2\pi i \mu_* J}}{\mu_* + \tilde{n}} - 2\pi i J E_1(-2\pi i J(\mu_* - \tilde{n})) + 2\pi i J E_1(2\pi i J(\mu_* + \tilde{n})) , \quad (\text{C.24})$$

where  $E_1(z) = \int_1^\infty dt e^{-zt}/t$  is the exponential integral with the large- $x$  expansion

$$E_1(ix) = e^{-ix} \left( \frac{1}{ix} + \frac{1}{x^2} + \mathcal{O}(x^{-3}) \right) . \quad (\text{C.25})$$

In the limit when  $\mu_* \gg \tilde{n}$ , the function  $h(J)$  then becomes

$$h(J) = \begin{cases} -\frac{2}{\mu_*} + \mathcal{O}(\mu_*^{-3}), & J = 0 , \\ \frac{2}{\mu_*} \frac{\sin(2\pi \mu_* J)}{2\pi \mu_* J} + \mathcal{O}(\mu_*^{-3}) , & J \neq 0 . \end{cases} \quad (\text{C.26})$$

The integral  $I_0(J)$  at a given value of  $J$  can then be evaluated by combining Eqs. (C.21), (C.22) and (C.26). For example,  $I_0(1)$  works out to be

$$\begin{aligned} I_0(1) &= \frac{1}{2} \left( h(1) - \frac{1}{2} h(0) - \frac{1}{2} h(2) \right) \\ &= \frac{1}{2} \left( \frac{\sin(2\pi \mu_*)}{\pi \mu_*^2} + \frac{1}{\mu_*} - \frac{1}{4} \frac{\sin(4\pi \mu_*)}{\pi \mu_*^2} \right) . \end{aligned} \quad (\text{C.27})$$

## C.2 Off-diagonal elements in mode-occupation number

When  $\Delta n \neq 0$ , the integral (C.17) can be split as

$$I_{\Delta n}(J) = \frac{1}{2} \left( g(J) - \frac{1}{2}g(J-1) - \frac{1}{2}g(J+1) \right) , \quad (\text{C.28})$$

where

$$g(J) = \frac{1}{\Delta n} \int_{-\mu}^{\mu_*} d\mu \left( \frac{e^{2\pi i J}}{\mu - \tilde{n}'} - (\tilde{n}' \leftrightarrow \tilde{n}) \right) . \quad (\text{C.29})$$

If  $J = 0$ , the integral can be computed exactly. For  $J \neq 0$ , it is possible to redefine the integration variables in the two summands so as to move the dependence on the expansion parameter  $\mu_*$  from the limits to the integrand:

$$g(J) = \frac{2 \cos(2\pi\mu_* J)}{\Delta n \mu_*} \int_{\tilde{n}'}^{\tilde{n}} dt \frac{e^{-2\pi i t J}}{1 - \left(\frac{t}{\mu_*}\right)^2} + \frac{2i \sin(2\pi\mu_* J)}{\Delta n \mu_*} \int_{\tilde{n}'}^{\tilde{n}} dt \frac{e^{-2\pi i t J}}{1 - \left(\frac{t}{\mu_*}\right)^2} \frac{t}{\mu_*} . \quad (\text{C.30})$$

One can then expand the integrand in the limit when  $\mu_* \gg \tilde{n}, \tilde{n}'$  and evaluate the integral to obtain

$$g(J) = \begin{cases} -\frac{2}{\mu_*} + \mathcal{O}(\mu_*^{-3}) , & J = 0 , \\ \frac{2}{\mu_*} \frac{\sin(2\pi\mu_* J)}{2\pi\mu_* J} + \mathcal{O}(\mu_*^{-3}) , & J \neq 0 . \end{cases} \quad (\text{C.31})$$

Comparing  $g(J)$  to the function  $h(J)$  in Eq. (C.26), we see that the two functions are equivalent,  $g(J) = h(J)$ . Hence, for  $J \neq 0$ , the two cases are coincident,

$$I_{\Delta n}(J) = I_0(J) , \quad J \neq 0 . \quad (\text{C.32})$$

On the other hand, at  $J = 0$ , the integral  $I_{\Delta n}$  differs from  $I_0(0)$  in Eq. (C.21) only by a zeroth-order, diagonal term:

$$\begin{aligned} I_{\Delta n}(0) &= \frac{1}{2} \left( h(0) - \frac{1}{2}h(-1) - \frac{1}{2}h(1) \right) \\ &= -\frac{1}{\mu_*} - \frac{\sin(2\pi\mu_*)}{2\pi\mu_*^2} + \mathcal{O}(\mu_*^{-3}) \\ &= I_0(0) - \pi^2 . \end{aligned} \quad (\text{C.33})$$

Hence, the corrections in both cases,  $\Delta n = 0$  and  $\Delta n \neq 0$ , coincide and are uniform in  $\Delta n$ .

## C.3 Final Result

Up to the leading order in  $\mu_*^{-1}$ , the integral (C.17) then becomes

$$I_{\Delta n}(J) = \pi^2 \delta_{\Delta n,0} \delta_{J,0} + \frac{1}{\mu_*} \left( -\delta_{J,0} + \frac{1}{2} \delta_{J,1} + \frac{1}{2} \delta_{J,-1} \right) + \mathcal{O}(\mu_*^{-2}) , \quad (\text{C.34})$$

where the second-order corrections are subleading for  $J = 0, \pm 1$ . In the other cases—i.e., when  $J \neq 0, \pm 1$ —it is the corrections at second order in  $\mu_*^{-1}$  that are dominant:

$$I_{\Delta n}(J) = \frac{1}{\mu_*} \left( \frac{\sin(2\pi\mu_* J)}{2\pi\mu_* J} - \frac{1}{2} \frac{\sin(2\pi\mu_*(J-1))}{2\pi\mu_*(J-1)} - \frac{1}{2} \frac{\sin(2\pi\mu_*(J+1))}{2\pi\mu_*(J+1)} \right) + \mathcal{O}(\mu_*^{-3}) . \quad (\text{C.35})$$

In order to interpret these results as corrections to the density-matrix elements, one has to employ the different definitions of  $J$  in Eq. (C.15) for the different cases.

### Interpretation

The matrix elements of the product  $(\bar{\beta}\beta)^{jn}_{j'n'}$  are proportional to

$$(\bar{\beta}\beta)^{jn}_{j'n'} \sim I_{\Delta n}(J_{\bar{\beta}\beta}) , \quad J_{\bar{\beta}\beta} = j' - j . \quad (\text{C.36})$$

Since the corrections are found to be uniform in  $\Delta n = n' - n$ , we can represent the matrix  $(\bar{\beta}\beta)^{jn}_{j'n'}$  as a tensor product of  $j$ 's and  $n$ 's. Then each matrix element that is labeled by  $j$  and  $j'$  is itself a uniform matrix in terms of  $n$  and  $n'$ . In this way, the product can be represented in the following form:

$$(\bar{\beta}\beta)^{jn}_{j'n'} \sim \begin{pmatrix} 1 & \frac{1}{\mu_*} & & \\ \frac{1}{\mu_*} & 1 & \frac{1}{\mu_*} & \frac{1}{\mu_*^2 J} \\ & \ddots & \ddots & \ddots \\ \frac{1}{\mu_*^2 J} & & \frac{1}{\mu_*} & 1 & \frac{1}{\mu_*} \\ & & \frac{1}{\mu_*} & & 1 \end{pmatrix} , \quad (\text{C.37})$$

where each depicted entry represents a uniform block and only the order of magnitude of the leading term is shown.

For the products  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$ , the corresponding index is  $J_{\alpha\beta} = \pm(j' + j + 1)$ , respectively. Since both  $j$  and  $j'$  are positive, these products have no zeroth-order contributions and only the matrix elements with  $j = j' = 0$  have corrections of the order  $\mathcal{O}(1/\mu_*)$ . All the other entries of the matrices  $\alpha\beta$  and  $\bar{\alpha}\bar{\beta}$  receive only second-order corrections  $\mathcal{O}(1/(\mu_*^2 J))$ .

How these findings impact upon the off-diagonal elements of the full density matrix is discussed in Subsection 5.1 of the main text.

## D Higher moments of the Marchenko–Pastur and $\rho_{SC}$ distributions

Here, we will quantify more precisely the differences between the Marchenko–Pastur (MP) distribution and the eigenvalue distribution of  $\rho_{SC}$ . This entails assigning the

distributions the same participation ratio and then determining how their higher moments are different.

Comparing their respective participation ratios (with  $n = N$ ) in Eqs. (5.17) and (5.23), one can readily identify the parameter  $c$  for the semiclassical two-point function,

$$c = \frac{1}{NC_{\text{BH}}(N)} . \quad (\text{D.1})$$

Let us consider the case  $NC_{\text{BH}} \gtrsim 1$ . This is really the regime of interest, since smaller values of  $NC_{\text{BH}}$  correspond to the case in which the Hawking model is valid (up to small corrections). We will further assume that the phases of the off-diagonal terms in  $\rho_{SC}$  can be treated as random. One can then estimate the higher moments of  $\rho_{SC}$  up to combinatorial factors and sub-leading terms in small  $1/NC_{\text{BH}}$ , which leads to

$$\frac{\text{tr}(\rho_{SC}^{2p})}{(\text{tr} \rho_{SC})^{2p}} \simeq \frac{N^{p+1} (C_{\text{BH}})^p}{N^{2p}} = \frac{(C_{\text{BH}})^p}{N^{p-1}} , \quad (\text{D.2})$$

$$\frac{\text{tr}(\rho_{SC}^{2p+1})}{(\text{tr} \rho_{SC})^{2p+1}} \simeq \frac{N^{p+1} (C_{\text{BH}})^p N}{N^{2p+1}} = \frac{(C_{\text{BH}})^p}{N^{p-1}} . \quad (\text{D.3})$$

The basic idea behind these estimates is that the off-diagonal parts of the matrices are dominant when  $NC_{\text{BH}} \gtrsim 1$  and the randomness of the phases requires these parts to sum coherently (*i.e.*, restricted to sums of the form  $\sum_{ij} M_{ij} M_{ji}$ ). A simple example should suffice to illustrate the point. Let  $\gamma$  and  $\eta$  be the diagonal and off-diagonal parts respectively of a matrix  $\rho$ . Applying the rule of coherent summation and, otherwise, insisting on the maximum power of  $\eta$ , we have, for the  $p = 2$  case,

$$\text{tr} \rho^4 = \sum_{ijkl} \rho_{ij} \rho_{jk} \rho_{kl} \rho_{li} \sim \sum_{ijl} \eta_{ij} \eta_{ji} \eta_{il} \eta_{li} , \quad (\text{D.4})$$

$$\text{tr} \rho^5 = \sum_{ijklm} \rho_{ij} \rho_{jk} \rho_{kl} \rho_{lm} \rho_{mi} \sim \sum_{ijl} \eta_{ij} \eta_{ji} \eta_{il} \eta_{li} \gamma_{ii} . \quad (\text{D.5})$$

As one can see, each of these traces results in  $3 = p + 1$  independent summations, a trend which continues for any value of  $p$ . This accounts for the factors of  $N^{p+1}$  in Eqs. (D.2) and (D.3); the rest is determined by the magnitude of the elements. These results indicate that the higher moments of the eigenvalue distribution are determined by an expansion in  $\sqrt{C_{\text{BH}}}$ .

The moments of the MP distribution, on the other hand, are expressed as a power series in  $c$  and not square roots thereof. Indeed, for the same conventions and the same regime of small  $c = 1/(NC_{\text{BH}})$ , the MP distribution would yield for the high moments

$$\frac{\text{tr}(\rho_{MP}^n)}{(\text{tr} \rho_{MP})^n} \simeq \left(\frac{1}{c}\right)^{n-1} \left(\frac{1}{N}\right)^{n-1} = (C_{\text{BH}})^{n-1} , \quad \text{for } n \gg 1 . \quad (\text{D.6})$$

This follows from the observation that the MP distribution (5.15) has, for small values of  $c$ , about  $Nc \sim 1/C_{\text{BH}}$  large and (roughly) equal-valued eigenvalues  $\lambda \sim 1/c$ .

This discrepancy between the MP distribution and the semiclassical distribution is a consequence of the square root of  $C_{\text{BH}}$  appearing in the off-diagonal elements of  $\rho_{SC}$ . Hence,  $\rho_{SC}$  does not precisely conform to an MP distribution nor should it necessarily be expected to. However, when  $NC_{\text{BH}} \sim 1$ , both expressions for the moments of the distributions scale in the same way,  $\frac{\text{tr}(\rho^n)}{(\text{tr } \rho)^n} \sim (C_{\text{BH}})^{n-1}$ . Therefore, we do expect that the two distributions share the same general features; in particular, once  $c = 1$  ( $NC_{\text{BH}} = 1$ ) is reached, both eigenvalue distributions begin to develop support for zero eigenvalues.

Yet, when  $NC_{\text{BH}} \gg 1$  is true, the higher moments of the eigenvalue distribution of  $\rho_{SC}$  are much more suppressed than those of the *MP*-distribution. For instance, the  $n^{\text{th}}$  semiclassical moment is smaller by a relative factor  $1/(NC_{\text{BH}})^{n/2}$  than its MP counterpart. Nevertheless, we do expect that, in this case, both distributions have only a few large eigenvalues but apparently differ in the detail. It would be interesting to find out what is the actual eigenvalue distribution of the semiclassical matrix.

## References

- [1] S. W. Hawking, “Particle Creation by Black Holes,” *Commun. Math. Phys.* **43** (1975) 199 [Erratum-*ibid.* **46** (1976) 206].
- [2] S. W. Hawking, “Breakdown of Predictability in Gravitational Collapse,” *Phys. Rev. D* **14** (1976) 2460.
- [3] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, “Black Holes: Complementarity or Firewalls?,” *JHEP* **1302**, 062 (2013) [arXiv:1207.3123 [hep-th]].
- [4] S. L. Braunstein, S. Pirandola and K. Zyczkowski, “Entangled black holes as ciphers of hidden information,” *Physical Review Letters* 110, **101301** (2013) [arXiv:0907.1190 [quant-ph]].
- [5] N. Itzhaki, “Is the black hole complementarity principle really necessary?,” arXiv:hep-th/9607028.
- [6] S. D. Mathur, “What Exactly is the Information Paradox?,” *Lect. Notes Phys.* **769**, 3 (2009) [arXiv:0803.2030 [hep-th]]; “The Information paradox: A Pedagogical introduction,” *Class. Quant. Grav.* **26**, 224001 (2009) [arXiv:0909.1038 [hep-th]]; “What does strong subadditivity tell us about black holes?,” *Nucl. Phys. Proc. Suppl.* **251-252**, 16 (2014) [arXiv:1309.6583 [hep-th]].
- [7] D. Marolf and J. Polchinski, “Gauge/Gravity Duality and the Black Hole Interior,” *Phys. Rev. Lett.* **111**, 171301 (2013) [arXiv:1307.4706 [hep-th]].

- [8] R. Bousso, “Firewalls From Double Purity,” *Phys. Rev. D* **88**, 084035 (2013) [arXiv:1308.2665 [hep-th]]; “Frozen Vacuum,” *Phys. Rev. Lett.* **112**, 041102 (2014) [arXiv:1308.3697 [hep-th]].
- [9] S. W. Hawking, “Information loss in black holes,” *Phys. Rev. D* **72**, 084013 (2005) [arXiv:hep-th/0507171].
- [10] D. N. Page, “Average entropy of a subsystem,” *Phys. Rev. Lett.* **71**, 1291 (1993) [arXiv:gr-qc/9305007]; “Information in black hole radiation,” *Phys. Rev. Lett.* **71**, 3743 (1993) [arXiv:hep-th/9306083].
- [11] E. Lubkin, “Entropy of an  $n$ -system from its correlation with a  $k$ -reservoir,” *J. Math. Phys.* **19**, 1028 (1978).
- [12] D. Harlow, “Jerusalem Lectures on Black Holes and Quantum Information,” arXiv:1409.1231 [hep-th].
- [13] R. Brustein and A. J. M. Medved, “Restoring predictability in semiclassical gravitational collapse,” *JHEP* **1309** (2013) 015 [arXiv:1305.3139 [hep-th]].
- [14] R. Brustein and A. J. M. Medved, “Phases of information release during black hole evaporation,” *JHEP* **1402** (2014) 116 [arXiv:1310.5861 [hep-th], arXiv:1310.5861].
- [15] R. Brustein and M. Hadad, “Wave function of the quantum black hole,” *Phys. Lett. B* **718**, 653 (2012) [arXiv:1202.5273 [hep-th]].
- [16] R. Brustein, “Origin of the blackhole information paradox,” *Fortsch. Phys.* **62**, 255 (2014) [arXiv:1209.2686 [hep-th]].
- [17] R. Brustein and A. J. M. Medved, “How black holes burn,” arXiv:1407.4914 [hep-th] and to appear in *Physical Review D*.
- [18] R. Brustein and A. J. M. Medved, “Semiclassical black holes expose forbidden charges and censor divergent densities,” *JHEP* **1309**, 108 (2013) [arXiv:1302.6086 [hep-th]].
- [19] J. M. Maldacena, “Eternal black holes in anti-de Sitter,” *JHEP* **0304**, 021 (2003) [arXiv:hep-th/0106112].
- [20] J. L. F. Barbon and E. Rabinovici, “Geometry And Quantum Noise,” *Fortsch. Phys.* **62**, 626 (2014) [arXiv:1404.7085 [hep-th]].
- [21] G. Adesso, D. Girolami and A. Serafini, *Phys. Rev. Lett.* **109** (2012) 19, 190502 [arXiv:1203.5116 [quant-ph]].
- [22] D. N. Page, “Particle Emission Rates from a Black Hole: Massless Particles from an Uncharged, Nonrotating Hole,” *Phys. Rev. D* **13** (1976) 198.
- [23] G. Jaeger, *Quantum Information: An Overview*, Springer, Boston, 2007.
- [24] L. H. Ford, “Quantum field theory in curved space-time,” In \*Campos do Jordao 1997, Particles and fields\* 345-388 [arXiv:gr-qc/9707062].
- [25] S. D. Mathur, “The Information paradox: A Pedagogical introduction,” *Class. Quant. Grav.* **26** (2009) 224001 [arXiv:0909.1038 [hep-th]].

- [26] L. C. Malacarne, R. S. Mendes, and E. K. Lenzi, “Average entropy of a subsystem from its average Tsallis entropy,” *Phys. Rev. E* **65** 046131, (2002).
- [27] Hans-Jürgen Sommers and Karol Zyczkowski, “Statistical properties of random density matrices,” arXiv:quant-ph/0405031, *J. Phys. A* **37** (2004), 8457.
- [28] Karol Zyczkowski, Karol A. Penson, Ion Nechita, Benoit Collins, “Generating random density matrices,” arXiv:1010.3570 [quant-ph], *J. Math. Phys.* **52**, 062201 (2011).
- [29] V. A. Marchenko and L. A. Pastur, “Distribution of eigenvalues for some sets of random matrices”, *Mat. Sb. (N.S.)*, **72** (1967) 507.
- [30] K. Huang, *Introduction to Statistical Physics*, Taylor & Francis, London, 2001.
- [31] V. Meden, lecture notes on the quantum harmonic oscillator at finite temperature, <http://web.physik.rwth-aachen.de/~meden/vielteilchenneu/skriptka2.pdf>